

# Asymptotic Ferromagnetic Ordering of Energy Levels for the Heisenberg Model on Large Boxes

Bruno Nachtergaele<sup>1</sup>, Wolfgang Spitzer<sup>2</sup> and Shannon Starr<sup>3</sup>

<sup>1</sup> University of California at Davis, Davis, CA 95616, USA

<sup>2</sup> FernUniversität in Hagen, Fakultät für Mathematik und Informatik,  
LG Angewandte Stochastik, 58084 Hagen, Germany

<sup>3</sup> University of Alabama at Birmingham, Applied Mathematics, Birmingham, AL 35294-1170

2 September 2015

## Abstract

We prove a result for the spin-1/2 quantum Heisenberg ferromagnet on  $d$ -dimensional boxes  $\{1, \dots, L\}^d \subset \mathbb{Z}^d$ . For any  $n$ , if  $L$  is large enough, the Hamiltonian satisfies: among all vectors whose total spin is at most  $(L^d/2) - n$ , the minimum energy is attained by a vector whose total spin is exactly  $(L^d/2) - n$ .

**Keywords:** Heisenberg model, quantum spin systems, simple exclusion process, ordering of energy levels, Aldous ordering, spectral gap.

**MCS numbers:** 82B10, 81R05, 81R50.

## 1 Introduction

In this article, we prove an asymptotic result for the ferromagnetic Heisenberg model on boxes. We prove that the *ferromagnetic ordering of energy levels (FOEL)* property holds up to some level, if the box is large enough.

The *FOEL condition at level  $n$*  means the following. Consider a  $d$ -dimensional box  $\mathbb{B}^d(L)$ , defined to be  $\{1, \dots, L\}^d \subset \mathbb{Z}^d$ . The ferromagnetic Heisenberg Hamiltonian on  $\mathbb{B}^d(L)$  commutes with total spin, and the ground state subspace is the total spin subspace for the maximum possible spin  $s = \frac{1}{2}L^d$ . For any  $n$ , we may restrict the Hamiltonian to the subspace spanned by total spin eigenvectors whose total spin satisfies  $s \leq \frac{1}{2}L^d - n$ . Then the minimum energy eigenvalue among vectors in this subspace is attained by an eigenvector whose total spin is  $s = \frac{1}{2}L^d - n$ . Roughly stated: lower energies are attained at higher spin. We prove that, for fixed  $n$  and  $d$  in  $\{1, 2, \dots\}$ , there exists  $L_0(n, d) \in \{1, 2, \dots\}$  such that FOEL- $n$  holds on  $\mathbb{B}^d(L)$  for each  $L \geq L_0(n, d)$ .

## Discussion

This property was proposed in [24] as a ferromagnetic version of a famous theorem by Lieb and Mattis for antiferromagnets and ferrimagnets [20]. Moreover, the authors proved that FOEL holds at level  $n$  for all  $n \in \{0, 1, \dots\}$ , for any  $d = 1$  dimensional box/chain of sidelength  $L \geq 2n$ . This is an optimal result because total spin can never be less than 0. The proof from [24] did not extend to periodic boundary conditions. It relied on a special basis, called the “Hulthén bracket basis,” defined and studied in the context of Bethe ansatz solvable models by Temperley and Lieb [28]. Due to the basis considered by Temperley and Lieb, the result generalized to the XXZ model on 1-dimensional

boxes, which possesses symmetry of the quantum group  $\mathcal{U}_q(\mathfrak{sl}_2)$ . It also could be proved for higher spin  $SU(2)$  models, as in [26], as well as higher-spin  $\mathcal{U}_q(\mathfrak{sl}_2)$ -symmetric models, generalizing the XXZ model in 1-dimension. The latter was carried out in [23]. The FOEL property in 1d also has some applications, such as a detailed study of droplets in the thermodynamic limit for the XXZ chain [25].

In higher dimensions, or even for periodic boundary conditions, the argument based on the Hulthén bracket basis does not work. In fact, if  $L$  is even, and  $L > 4$ , and one considers the 1-dimensional torus of length  $L$  (i.e., the ring), then there is numerical evidence that FOEL is violated when one takes the level to be  $n = (L/2) - 1$ . See, for example, [27].

On the other hand, a deep and interesting result of Caputo, Liggett and Richthammer [5] showed that FOEL at level  $n = 1$  holds for all graphs. This had been conjectured, previously, by Aldous, with some evidence by Diaconis. Before [5], Handjani and Jungreis wrote an important paper including historical perspective [18]. Caputo, Liggett and Richthammer proved Aldous's conjecture. It is interesting to ask how one may generalize their results for  $n > 1$ , in a general way for all graphs? We do not address that question in this article, but there is important work in this direction [6, 1, 2, 13].

Before their major advance there were two closely related papers by Conomos and one of the authors [10] and Morris [22] essentially proving Aldous's conjecture for large boxes. These papers built on an earlier inductive argument idea of Koma and Nachtergaele [19], who had calculated the exact spectral gap for the XXZ model on 1-dimensional boxes/chains, and Handjani and Jungreis [18], who had considered examples of graphs for which Aldous's conjecture was provable at that time. The argument in [10] and [22] is easier than Caputo, Liggett and Richthammer's proof in [5]. In the present paper we generalize this argument from  $n = 1$  to  $n > 1$ . On large boxes with few spin deviations  $n$  a kinetic energy term forces the low energy states to "spread out." Then, for such states, the model is well approximated by an ideal Bose gas of "magnons."

Two years ago, a major advance was made by Correggi, Giuliani and Seiringer [11, 12]. They proved that the free energy at low temperature is asymptotically the same as for the non-interacting Bose gas on the lattice. One motivation for their work is the important open problem of proving  $SU(2)$  symmetry breaking at low temperature for the Heisenberg ferromagnet. See

<http://web.math.princeton.edu/~aizenman/OpenProblems.iamp/9901.HeisenbergFerr.html>.

There is a relation between our result and theirs. A lemma in their paper establishes that the minimum energy among all eigenvectors with total spin  $s = \frac{1}{2}L^d - n$  is at least  $Cn$  for some  $C > 0$ , uniformly in  $L$  and  $n$  (relative to the ground state energy 0). Their constant  $C$  is not sharp for small  $n$ , but their bound is uniform in  $n$ . We will obtain a sharp bound for small  $n$ , but not uniformly.

## 2 Set-up and Statement of the main result

Consider the Heisenberg model on a finite graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with vertex set  $\mathcal{V}$  and edge set  $\mathcal{E}$ . We denote a reference Hilbert space  $\mathcal{H} \cong \mathbb{C}^2$  for a single spin with orthonormal basis  $|\uparrow\rangle, |\downarrow\rangle$ . Relative to this basis, the Pauli spin matrices  $\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}$ , and raising/lowering operators are

$$\sigma^{(1)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma^{(2)} = i \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \sigma^{(3)} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \sigma^+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \sigma^- = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

The spin-1/2 spin matrices are  $S^{(a)} = \frac{1}{2}\sigma^{(a)}$ , for  $a \in \{1, 2, 3\}$ , and  $S^\pm = \sigma^\pm = S^{(1)} \pm iS^{(2)}$ .

**Remark 2.1** For each  $a \in \{1, 2, 3\}$ , we have  $(\sigma^{(a)})^* = \sigma^{(a)}$ . Also,  $(\sigma^+)^* = \sigma^-$ .

For each  $x \in \mathcal{V}$ , there is an isomorphic copy  $\mathcal{H}_x \cong \mathcal{H}$ . The total Hilbert space for the Heisenberg model on  $\mathcal{V}$  is  $\mathcal{H}_{\mathcal{V}} = \bigotimes_{x \in \mathcal{V}} \mathcal{H}_x$ . For each  $x \in \mathcal{V}$ , and  $a \in \{1, 2, 3\}$ , we let  $S_x^{(a)}$  denote the operator on  $\mathcal{H}_{\mathcal{V}}$  such that, for simple tensor product vectors,  $\bigotimes_{y \in \mathcal{V}} \psi_y$  (where  $\psi_y \in \mathcal{H}_y$  for each  $y \in \mathcal{V}$ ),

$$S_x^{(a)} \bigotimes_{y \in \mathcal{V}} \psi_y = \bigotimes_{y \in \mathcal{V}} \psi'_y \quad \text{for} \quad \psi'_y = \begin{cases} S_x^{(a)} \psi_x & \text{for } y = x, \\ \psi_y & \text{for } y \in \mathcal{V} \setminus \{x\}. \end{cases}$$

We define  $S_x^{\pm} = S_x^{(1)} \pm iS_x^{(2)}$ , consistently. The Heisenberg Hamiltonian  $H_{\mathcal{G}}$  is the operator on  $\mathcal{H}_{\mathcal{V}}$ :

$$H_{\mathcal{G}} = \sum_{\{x,y\} \in \mathcal{E}} h_{xy}, \quad \text{where} \quad h_{xy} = \frac{1}{4} \mathbb{1} - \sum_{a=1}^3 S_x^{(a)} S_y^{(a)} = \frac{1}{4} \mathbb{1} - S_x^{(3)} S_y^{(3)} - \frac{1}{2} S_x^+ S_y^- - \frac{1}{2} S_x^- S_y^+, \quad (1)$$

where  $\mathbb{1}$  denotes the identity operator on  $\mathcal{H}_{\mathcal{V}}$ . The model is  $\text{SU}(2)$  invariant. We may define

$$S_{\mathcal{V}}^{(a)} = \sum_{x \in \mathcal{V}} S_x^{(a)}, \quad \text{for } a = 1, 2, 3, \quad \text{and} \quad S_{\mathcal{V}}^{\pm} = \sum_{x \in \mathcal{V}} S_x^{\pm} = S_{\mathcal{V}}^{(1)} \pm iS_{\mathcal{V}}^{(2)}, \quad (2)$$

Then  $[H_{\mathcal{G}}, S_{\mathcal{V}}^{(a)}] = 0$  for  $a \in \{1, 2, 3\}$  which also implies  $[H_{\mathcal{G}}, S_{\mathcal{V}}^{\pm}] = 0$ .

We recall basic  $\text{SU}(2)$  theory to state the main result. By Remark 2.1 and the definitions  $(S_{\mathcal{V}}^{(a)})^* = S_{\mathcal{V}}^{(a)}$  for  $a \in \{1, 2, 3\}$ , and  $(S_{\mathcal{V}}^+)^* = S_{\mathcal{V}}^-$ . Let  $\Omega_{\mathcal{V}}(\emptyset) \in \mathcal{H}_{\mathcal{V}}$  be the simple tensor product vector, which is the product of  $|\uparrow\rangle \in \mathcal{H}_x$ , for each  $x \in \mathcal{V}$ . Hence,  $\Omega_{\mathcal{V}}(\emptyset) \in \ker(S_x^{(3)} - \frac{1}{2}\mathbb{1})$  for each  $x \in \mathcal{V}$ . For each  $n \in \{1, 2, \dots\}$  and each  $(x_1, \dots, x_n) \in \mathcal{V}^n$ , let

$$\Omega_{\mathcal{V}}^{(n)}(x_1, \dots, x_n) \stackrel{\text{def}}{=} S_{x_1}^- \cdots S_{x_n}^- \Omega_{\mathcal{V}}(\emptyset). \quad (3)$$

Then, for any  $\pi \in S_n$ , we have  $\Omega_{\mathcal{V}}^{(n)}(x_{\pi_1}, \dots, x_{\pi_n}) = \Omega_{\mathcal{V}}^{(n)}(x_1, \dots, x_n)$  since  $S_{x_1}^-, \dots, S_{x_n}^-$  commute. Given any subset  $X \subseteq \mathcal{V}$ , let us define  $\Omega_{\mathcal{V}}(X)$  as follows: if  $|X| = n$  for some  $n \in \{1, \dots, |\mathcal{V}|\}$  and  $X = \{x_1, \dots, x_n\}$ , then we define  $\Omega_{\mathcal{V}}(X)$  to be  $\Omega_{\mathcal{V}}^{(n)}(x_1, \dots, x_n)$ . Then we may easily deduce:

- For  $x, y \in \mathcal{V}$ ,  $[S_x^{(3)}, S_y^{\pm}] = \pm \delta_{x,y} S_y^{\pm}$  and  $[S_x^+, S_y^-] = 2\delta_{x,y} S_x^{(3)}$ .
- Summing,  $[S_{\mathcal{V}}^{(3)}, S_x^{\pm}] = \pm S_x^{\pm}$  for each  $x \in \mathcal{V}$ . Also,  $[S_{\mathcal{V}}^{(3)}, S_{\mathcal{V}}^{\pm}] = \pm S_{\mathcal{V}}^{\pm}$  and  $[S_{\mathcal{V}}^+, S_{\mathcal{V}}^-] = 2S_{\mathcal{V}}^{(3)}$ .
- For  $X \subseteq \mathcal{V}$  and  $x \in \mathcal{V}$ ,  $S_x^+ \Omega_{\mathcal{V}}(X) = \mathbf{1}_X(x) \Omega_{\mathcal{V}}(X \setminus \{x\})$  and  $S_x^- \Omega_{\mathcal{V}}(X) = \mathbf{1}_{\mathcal{V} \setminus X}(x) \Omega_{\mathcal{V}}(X \cup \{x\})$ .

For each  $m \in \mathbb{R}$ , we define the  $m$ -magnon subspace,

$$\mathcal{H}_{\mathcal{V}}^{\text{mag}}(m) \stackrel{\text{def}}{=} \ker(S_{\mathcal{V}}^{(3)} - M \cdot \mathbb{1}), \quad \text{for } M = \frac{1}{2}|\mathcal{V}| - m. \quad (4)$$

We denote the *Casimir* operator by the symbol  $\mathcal{C}_{\mathcal{V}} = \sum_{a=1}^3 (S_{\mathcal{V}}^{(a)})^2 = (S_{\mathcal{V}}^{(3)})^2 + \frac{1}{2} S_{\mathcal{V}}^+ S_{\mathcal{V}}^- + \frac{1}{2} S_{\mathcal{V}}^- S_{\mathcal{V}}^+$ . This is also called the total-spin operator,  $\mathbf{S}_{\mathcal{V}}^2$ . For  $n \in \mathbb{R}$ , we define the  $n$ -spin deviate subspace

$$\mathcal{H}_{\mathcal{V}}^{\text{spin}}(n) = \ker(\mathcal{C}_{\mathcal{V}} - s(s+1)\mathbb{1}), \quad \text{for } s = \frac{1}{2}|\mathcal{V}| - n. \quad (5)$$

A calculation shows  $[\mathcal{C}_{\mathcal{V}}, S_{\mathcal{V}}^{(a)}] = 0$  for  $a \in \{1, 2, 3\}$ . So we can consider simultaneous eigenspaces:

$$\mathcal{H}_{\mathcal{V}} = \bigoplus_{(m,n) \in \mathcal{A}_{\mathcal{V}}} \mathcal{H}_{\mathcal{V}}(m, n), \quad \text{where}$$

$$\forall (m, n) \in \mathbb{R}^2, \quad \mathcal{H}_{\mathcal{V}}(m, n) \stackrel{\text{def}}{=} \mathcal{H}_{\mathcal{V}}^{\text{mag}}(m) \cap \mathcal{H}_{\mathcal{V}}^{\text{spin}}(n), \quad \text{and}$$

$$\mathcal{A}_{\mathcal{V}} \stackrel{\text{def}}{=} \{(m, n) \in \mathbb{R}^2 : \mathcal{H}_{\mathcal{V}}(m, n) \neq \{0\}\}.$$

Since  $\mathcal{C}_\mathcal{V}$  is in the subalgebra generated by  $S_\mathcal{V}^{(a)}$ , for  $a \in \{1, 2, 3\}$ , we see that  $H_\mathcal{G}$  also commutes with  $\mathcal{C}_\mathcal{V}$ . Hence,  $\mathcal{H}_\mathcal{V}(m, n)$  is an invariant subspace for  $H_\mathcal{G}$  for each  $(m, n) \in \mathcal{A}_\mathcal{V}$ . Also note that since  $[\mathcal{C}_\mathcal{V}, S_\mathcal{V}^{(a)}] = 0$  for  $a \in \{1, 2, 3\}$ , then  $[\mathcal{C}_\mathcal{V}, S_\mathcal{V}^\pm] = 0$ , by 2). The following lemma gives the structure of the  $\text{SU}(2)$  representation on  $\mathcal{H}_\mathcal{V}$ . We will not prove this lemma, which is standard. (See, e.g., [16].)

Let  $\mathcal{P}(\mathcal{V})$  denote the power set of  $\mathcal{V}$  and let

$$\forall m \in \{0, \dots, |\mathcal{V}|\}, \quad \mathcal{P}_m(\mathcal{V}) \stackrel{\text{def}}{=} \{X \in \mathcal{P}(\mathcal{V}) : |X| = m\}. \quad (6)$$

**Lemma 2.2** (1) For  $m \in \{0, \dots, |\mathcal{V}|\}$ , we have  $\mathcal{H}_\mathcal{V}^{\text{mag}}(m) = \text{span}(\{\Omega_\mathcal{V}(X) : X \in \mathcal{P}_m(\mathcal{V})\})$ , and  $\langle \Omega_\mathcal{V}(X), \Omega_\mathcal{V}(Y) \rangle = \delta_{X,Y}$  for  $X, Y \in \mathcal{P}_m(\mathcal{V})$ . In particular,  $\dim(\mathcal{H}_\mathcal{V}^{\text{mag}}(m)) = \binom{|\mathcal{V}|}{m}$ .  
(2) For  $m \in \{0, \dots, |\mathcal{V}| - 1\}$ , we have  $S_\mathcal{V}^- \mathcal{H}_\mathcal{V}^{\text{mag}}(m) \subseteq \mathcal{H}_\mathcal{V}^{\text{mag}}(m+1)$ , and  $S_\mathcal{V}^- \mathcal{H}_\mathcal{V}^{\text{mag}}(|\mathcal{V}|) = 0$ .  
(3) Define  $\mathcal{A}'(\mathcal{V})$  to be  $\{(m, n) \in \mathbb{Z}^2 : m \geq 0, n \geq 0, n \leq m, m+n \leq |\mathcal{V}|\}$ . For each  $(m, n) \in \mathcal{A}'(\mathcal{V})$ , define  $\mathcal{H}'_\mathcal{V}(m, n) \subset \mathcal{H}_\mathcal{V}$  as  $\mathcal{H}'_\mathcal{V}(m, n) = (S_\mathcal{V}^-)^{m-n} (\mathcal{H}_\mathcal{V}^{\text{mag}}(n) \cap \ker(S_\mathcal{V}^+))$ , where  $(S_\mathcal{V}^-)^0 = \mathbb{1}$ . Then  $\mathcal{H}'_\mathcal{V}(m, n) \subseteq \mathcal{H}_\mathcal{V}(m, n)$ .  
(4) For each  $(m, n) \in \mathcal{A}'(\mathcal{V})$ , we have  $\mathcal{H}'_\mathcal{V}(m, n) \subseteq \ker(S_\mathcal{V}^+ S_\mathcal{V}^- - (s+M)(s-M+1)\mathbb{1})$  where  $s = \frac{1}{2}|\mathcal{V}| - n$ ,  $M = \frac{1}{2}|\mathcal{V}| - m$ .  
(5) For each  $(m, n) \in \mathcal{A}'(\mathcal{V})$ , we have  $\dim(\mathcal{H}'_\mathcal{V}(m, n)) = \binom{|\mathcal{V}|}{n} - \binom{|\mathcal{V}|}{n-1}$ , where  $\binom{|\mathcal{V}|}{-1}$  is set to 0.  
(6) The combinatorial identity  $\sum_{(m,n) \in \mathcal{A}'(\mathcal{V})} \dim(\mathcal{H}'_\mathcal{V}(m, n)) = \dim(\mathcal{H}_\mathcal{V})$  is true. So  $\mathcal{A}(\mathcal{V})$  is  $\mathcal{A}'(\mathcal{V})$ . And, for each  $(m, n) \in \mathcal{A}(\mathcal{V})$ , we have that  $\mathcal{H}_\mathcal{V}(m, n)$  is  $\mathcal{H}'_\mathcal{V}(m, n)$ .

The following is key for us.

**Proposition 2.3**  $\min \text{spec}(H_\mathcal{G}) = 0$ . Moreover, if  $\mathcal{G}$  is connected, then  $\ker(H_\mathcal{G}) = \mathcal{H}_\mathcal{V}^{\text{spin}}(0)$ .

Proposition 2.3 is one small part of a famous result of Lieb and Mattis on “ordering of energy levels,” [20]. We will not reprove Proposition 2.3. But the reader is strongly recommended to consult [20] if not already acquainted with that article. It provides the main motivation for the present article.

**Corollary 2.4** For any finite graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , we have that  $H_\mathcal{G}$  is positive semi-definite.

**Proof:** From Proposition 2.3 all eigenvalues are nonnegative. □

For operators,  $A \geq B$  means  $A - B$  is positive semi-definite. Hence, for later reference, we record:

$$H_\mathcal{G} \geq 0. \quad (7)$$

We may denote by  $\ell^2(\mathcal{V})$  the finite dimensional Hilbert space of all functions  $F : \mathcal{V} \rightarrow \mathbb{C}$  with  $\|F\|^2 = \sum_{x \in \mathcal{V}} |F(x)|^2$ . Then the graph Laplacian is an operator  $-\Delta_\mathcal{G}$  on  $\ell^2(\mathcal{V})$ , given by the formula

$$-\Delta_\mathcal{G} F(x) = \frac{1}{2} \sum_{y \in \mathcal{N}(\mathcal{E}, x)} (F(x) - F(y)), \quad (8)$$

for each  $x \in \mathcal{V}$ , where we define the neighborhood of  $x$  as

$$\mathcal{N}(\mathcal{E}, x) \stackrel{\text{def}}{=} \{y \in \mathcal{V} : (x, y) \in \mathcal{E}\}. \quad (9)$$

Then, for any  $F \in \ell^2(\mathcal{V})$ ,

$$\langle F, -\Delta_\mathcal{G} F \rangle = \frac{1}{2} \sum_{\{x, y\} \in \mathcal{E}} |F(x) - F(y)|^2. \quad (10)$$

This is related to (7) as it seems was first observed by Toth [30], because for each  $n \in \{0, \dots, |\mathcal{V}|\}$  there is a graph derived from  $\mathcal{G}$  such that  $H_{\mathcal{G}} \upharpoonright \mathcal{H}^{\text{mag}}(n)$  is unitarily equivalent to the Laplacian for that graph, if one restricts to symmetric tensors, as one does for a Bose gas.

Let us now explain the ferromagnetic ordering of energy levels property, which we call FOEL, for short. For each  $n \in \{0, \dots, \lfloor \frac{1}{2}|\mathcal{V}| \rfloor\}$ , we define  $\mathfrak{E}_n(\mathcal{G})$  as the minimum energy among  $n$ -spin deviates:

$$\mathfrak{E}_n(\mathcal{G}) = \min \text{spec} (H_{\mathcal{G}} \upharpoonright \mathcal{H}_{\mathcal{V}}^{\text{spin}}(n)) = \min \left\{ \lambda \in \mathbb{R} : \mathcal{H}_{\mathcal{V}}^{\text{spin}}(n) \cap \ker(H_{\mathcal{G}} - \lambda \mathbb{1}) \neq \{0\} \right\}. \quad (11)$$

Then Proposition 2.3 may be restated as follows:  $0 = \mathfrak{E}_0(\mathcal{G}) \leq \mathfrak{E}_n(\mathcal{G})$  for all  $n \geq 0$ , and there is a strict inequality for all  $n > 0$  if  $\mathcal{G}$  is connected.

**Definition 2.5** *Given  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  and  $n \in \{0, \dots, \lfloor \frac{1}{2}|\mathcal{V}| \rfloor\}$ , we say  $\mathcal{G}$  satisfies FOEL- $n$  (strict FOEL- $n$ ) if  $\mathfrak{E}_{n'}(\mathcal{G}) \geq \mathfrak{E}_n(\mathcal{G})$  for all  $n' \geq n$  ( $\mathfrak{E}_{n'}(\mathcal{G}) > \mathfrak{E}_n(\mathcal{G})$  for any  $n' > n$ ).*

By Proposition 2.3 every graph  $\mathcal{G}$  satisfies FOEL-0, and it satisfies strict FOEL-0 if  $\mathcal{G}$  is connected. The FOEL property was introduced in [24] as a ferromagnetic analogue of Lieb and Mattis's "ordering of energy levels" property [20], which they proved for antiferromagnets and ferrimagnets on bipartite lattices.

It is of interest to verify FOEL for a collection of graphs arising in physical models. We note that  $\mathbb{Z}^d$  has a natural graph structure. Denoting an element of  $\mathbb{Z}^d$  as  $\mathbf{r} = (r_1, \dots, r_d)$ , the graph distance to the origin corresponds to the  $\ell^1$ -norm  $\|\mathbf{r}\|_1 = |r_1| + \dots + |r_d|$ . Setting  $\mathcal{V} = \mathbb{Z}^d$ , then  $\{\mathbf{r}, \mathbf{r}'\}$  is an edge in  $\mathcal{E}(\mathbb{Z}^d)$  for  $\mathbf{r}, \mathbf{r}' \in \mathbb{Z}^d$  if and only if  $\|\mathbf{r} - \mathbf{r}'\|_1 = 1$ . Given any subset  $\Lambda \subseteq \mathbb{Z}^d$ , we will write  $\mathcal{G}(\Lambda)$  for the graph where  $\mathcal{V}(\Lambda) = \Lambda$  and  $\mathcal{E}(\Lambda)$  is the edge set induced from  $\mathbb{Z}^d$ : in other words,

$$\mathcal{E}(\Lambda) = \{\{\mathbf{r}, \mathbf{r}'\} \subseteq \Lambda : \{\mathbf{r}, \mathbf{r}'\} \in \mathcal{E}(\mathbb{Z}^d)\}. \quad (12)$$

When  $\Lambda \subset \mathbb{Z}^d$  is finite, we will write  $H_{\Lambda}$  in place of  $H_{\mathcal{G}(\Lambda)}$  when no confusion arises from this notation. Similarly, we write  $\mathfrak{E}_n(\Lambda)$  in place of  $\mathfrak{E}_n(\mathcal{G}(\Lambda))$ . We will say that  $\Lambda$  satisfies (strict) FOEL- $n$  if  $\mathcal{G}(\Lambda)$  does.

For each  $d$  and  $L$  in  $\{1, 2, \dots\}$  we define the  $d$ -dimensional box of sidelength  $L$

$$\mathbb{B}^d(L) = \{\mathbf{r} = (r_1, \dots, r_d) \in \mathbb{Z}^d : r_1, \dots, r_d \in \{1, \dots, L\}\}.$$

This is a subset of  $\mathbb{Z}^d$ . We will prove the FOEL properties for boxes.

**Theorem 2.6 (Main Result)** *For each choice of  $n, d \in \{1, 2, \dots\}$ , there exists a finite integer  $L_{n,d}^0 \in \{1, 2, \dots\}$  such that  $\Lambda = \mathbb{B}^d(L)$  satisfies strict FOEL- $n$  for all  $L \geq L_{n,d}^0$ .*

In Figure 1 we have plotted the numerical eigenvalues of  $H_{\mathcal{G}}$  against the spin  $s = \frac{1}{2}|\mathcal{V}| - n$ , for two examples of graphs: a chain  $\{1, \dots, 8\}$  (i.e.,  $B^1(8)$ ) viewed as a subgraph of  $\mathbb{Z}$ , and a  $3 \times 3$  box (i.e.,  $B^2(3)$ ) viewed as a subgraph of  $\mathbb{Z}^2$ . The figure shows that they both satisfy FOEL- $n$  for all  $n$ . These are both boxes, one of dimension  $d = 1$  and one of dimension  $d = 2$ . Our theorem applies to sufficiently large boxes.

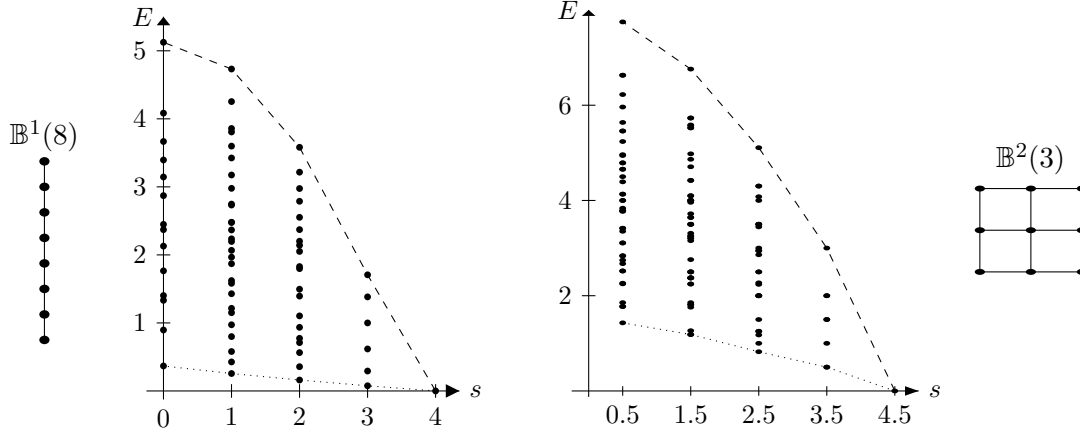


Figure 1: We have plotted the eigenvalues for  $H_{\mathbb{B}^d(L)}$  for two graphs. On the left is  $\mathbb{B}^1(8)$ . On the right is  $\mathbb{B}^2(3)$ . (The graphs are indicated on the sides.) On the horizontal axes, we have plotted the spin parameter  $s = \frac{1}{2}L^d - n$ . In each vertical column above this, we have plotted all the energy eigenvalues  $E \in \text{spec}(H_{\mathbb{B}^d(L)} \upharpoonright \mathcal{H}_{\mathbb{B}^d(L)}^{\text{spin}}(n))$ . The lowest energy eigenvalue in each column is  $\mathfrak{E}_n(\mathbb{B}^d(L))$ . We have plotted a dotted line to aid the eye in comparing these values for different values of  $s = \frac{1}{2}L^d - n$ . In both of these graphs  $\mathfrak{E}_n(\mathbb{B}^d(L))$  is an increasing function of  $n$  (decreasing function of  $s$ ). That property is FOEL. A related property is Lieb and Mattis's antiferromagnetic ordering of energy levels. The Lieb-Mattis line is dashed.

## 2.1 Outline for the rest of the paper

In Section 3, we state the two key steps for the proof of Theorem 2.6: an inductive argument for proving FOEL, and an approximate analysis of low energy eigenvectors of the Hamiltonian. In Section 4, we then give the conditional proof of the main theorem, conditional on the two key steps. The rest of the paper is dedicated to proving the two key steps.

In Section 5, we prove the first key step, which is an inductive argument for the proof of FOEL. It relies on some important information for the low energy eigenvectors of the Hamiltonian: not just one one graph, but on an increasing sequence of graphs. This is what the second key step is devoted to proving. In Section ?? we prove the second key step. This is a partial verification of the linear spin wave approximation, at very low energies. In turn the proof in Section ?? relies on two discrete analogues of Sobolev inequalities, which are proved in the following two sections. The first discrete Sobolev inequality is a trace theorem type bound, proved in Section ?. The second discrete Sobolev inequality is an extension theorem, which is proved in ?. A number of elementary results are relegated to appendices, which appear at the end of the paper.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Set-up and Statement of the main result</b>	<b>2</b>
2.1	Outline for the rest of the paper . . . . .	6
<b>3</b>	<b>The two key steps of the proof</b>	<b>7</b>
3.1	Step I: An Inductive Argument for Establishing FOEL . . . . .	7
3.2	Step II: Analysis of low energy wave-functions . . . . .	8

4	Conditional Proof of the Main Theorem	11
5	Proof of Key Step I: Inductive argument for FOEL	14
6	Proof of Key Step II – Part A: Linear Spin Wave Approximation	18
7	Proof of Key Step II – Part B: Variational Argument	21
8	Proof of Key Step II – Part C: Trace theorem type bound	24
9	Proof of Key Step II – Part D: Extension theorem type bound	27
A	Proof of advanced calculus fact for “new lows”	33
B	Proof of summary of spectrum of $\mathfrak{H}_{\mathbb{B}^d(L)}^{(n)}$	34
C	Proofs of easy facts about the graph Laplacians	34
D	Proof of Proposition 3.9	34

### 3 The two key steps of the proof

The proof of Theorem 2.6 relies on two main ideas. The first key idea is an inductive argument for FOEL. To obtain the induction step for that argument, we need some asymptotic approximation to the low energy spectrum. The second main idea is a method to obtain that asymptotic approximation. It is a mathematical verification of part of the physicists’ *linear spin wave approximation*. This follows by using some discrete versions of Sobolev type inequalities to show that states with low energy are *spread out*, which seems to be a key physical assumption.

#### 3.1 Step I: An Inductive Argument for Establishing FOEL

The main idea for an inductive proof of FOEL is the following.

**Proposition 3.1** *Suppose that  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  is a finite graph, and that  $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$  is a graph such that  $\mathcal{V}' = \mathcal{V} \cup \{x'\}$  for a single vertex  $x'$  not in  $\mathcal{V}$ , and such that  $\mathcal{E} \subseteq \mathcal{E}'$ . Suppose that  $\mathcal{G}$  satisfies FOEL- $n$ , for some  $n \in \{0, 1, \dots, \lfloor \frac{1}{2}|\mathcal{V}| \rfloor\}$ . Suppose, further that the following condition is satisfied*

$$\mathfrak{E}_n(\mathcal{G}') \leq \mathfrak{E}_n(\mathcal{G}).$$

*Then  $\mathcal{G}'$  also satisfies FOEL- $n$ .*

This proposition was originally proved in [19] for the special case of  $n = 1$ . In [24], the present authors gave the argument generalizing the proof to higher  $n$ . In Section 5 we will state and prove a generalization of this fact. The argument from [19] was also, independently, rediscovered in [18], slightly later.

**Corollary 3.2** *For any  $n \in \mathbb{N}$  and for any  $N \in \{n+1, n+2, \dots\}$ , suppose that  $\mathcal{G}_{2n}, \mathcal{G}_{2n+1}, \dots, \mathcal{G}_N$  forms a family of graphs, satisfying the following conditions.*

(i) *For each  $k \in \{2n, \dots, N\}$ , we have that  $\mathcal{G}_k = (\mathcal{V}_k, \mathcal{E}_k)$  satisfies  $|\mathcal{V}_k| = k$ .*

(ii) *For each  $k \in \{2n, \dots, N-1\}$ , we have that*

$$(a) \quad \mathcal{V}_k \subset \mathcal{V}_{k+1},$$

- (b)  $\mathcal{E}_k \subset \mathcal{E}_{k+1}$ ,
- (c)  $|\mathcal{V}_{k+1} \setminus \mathcal{V}_k| = 1$ .

(iii) For each  $k \in \{2n, \dots, N-1\}$ , we have that  $\mathfrak{E}_n(\mathcal{G}_{k+1}) \leq \mathfrak{E}_n(\mathcal{G}_k)$ .

Then  $\mathcal{G}_k$  satisfies FOEL- $n$  for each  $k \in \{2n, \dots, N\}$ .

**Proof:** According to Definition 2.5, the graph  $\mathcal{G}_{2n}$  trivially satisfies FOEL- $n$ , since  $\frac{1}{2}|\mathcal{V}_n| = n$  so that the only choice for  $n'$  greater than or equal to  $n$  (and less than or equal to  $\frac{1}{2}|\mathcal{V}_n|$ ) is  $n$ , itself. But then, using Proposition 3.1, we may establish that  $\mathcal{G}_k$  satisfies FOEL- $n$  for all  $k \in \{2n, \dots, N\}$ , with  $k = 2n$  just established as the initial step.  $\square$

This corollary was used in [24, 26] and [23] to prove FOEL- $n$  for all  $n$  in one-dimensional models. Let us say that graphs  $\mathcal{G}_{2n}, \mathcal{G}_{2n+1}, \dots, \mathcal{G}_N$  satisfying conditions (i) and (ii) from the hypotheses of Corollary 3.2 are a “growing family of graphs.”

**Proposition 3.3** *Suppose  $n \in \mathbb{N}$  and  $N \in \{n+1, n+2, \dots\}$  are fixed. Suppose  $\mathcal{G}_{2n}, \mathcal{G}_{2n+1}, \dots, \mathcal{G}_N$  are a growing family of graphs. In place of condition (iii), suppose that the following holds:*

(iii')  $\min\{\mathfrak{E}_n(\mathcal{G}_k) : k \in \{2n, \dots, N\}\} = \mathfrak{E}_n(\mathcal{G}_N)$ .

Then

$$\mathfrak{E}_n(\mathcal{G}_N) \leq \min\{\mathfrak{E}_r(\mathcal{G}_k) : k \in \{2n, \dots, N\}, r \in \{n, \dots, \lfloor \frac{1}{2}|\mathcal{V}_k| \rfloor\}\}. \quad (13)$$

In particular,  $\mathcal{G}_N$  satisfies FOEL- $n$ .

We will prove this result in Section 5. For the special case of  $n = 1$ , the analogous result was proved in [10]. It was also, independently, and slightly later, proved in [22] for the special case of  $n = 1$ . The present extension to  $n > 1$ , follows immediately by the same argument. A key part of the idea is to introduce coupling constants (known as “rates” in the probability and stochastic processes community) for the purpose of diluting the graphs. The idea of including rates in such arguments can be traced back at least to Handjani and Jungreis [18].

The hypothesis of Proposition 3.3 is weaker than that of Corollary 3.2 because the condition (iii') is weaker than the condition (iii). But the conclusion is also weaker since we only conclude FOEL- $n$  for the final graph  $\mathcal{G}_N$ , as opposed to all graphs  $\mathcal{G}_k$ , for  $k \in \{2n, \dots, N\}$ . For a certain family of graphs, verifying condition (iii') is easier than trying to verify condition (iii). Let us define the family of graphs, now.

**Definition 3.4** *Given  $d \in \mathbb{N}$  and for each  $N \in \mathbb{N}$ , define  $L(d, N) = \lfloor N^{1/d} \rfloor$  and  $L^+(d, N) = \lceil N^{1/d} \rceil$ . Let  $\Lambda(d, N) \subset \mathbb{Z}^d$  be defined as follows. If  $L^+(d, N) = L(d, N)$ , then  $\Lambda(d, N) \stackrel{\text{def}}{=} \mathbb{B}_{L(d, N)}^d$ . Otherwise, let  $\prec_d$  denote the lexicographic ordering on  $\mathbb{Z}^d$ , and let  $\Lambda(d, N) \stackrel{\text{def}}{=} \mathbb{B}_{L(d, N)}^d \cup S(d, N)$  where  $S(d, N)$  is the set of cardinality  $N - (L(d, N))^d$ , consisting of points in  $\mathbb{B}_{L^+(d, N)}^d \setminus \mathbb{B}_{L(d, N)}^d$  which are smallest with respect to  $\prec_d$ .*

### 3.2 Step II: Analysis of low energy wave-functions

We continue to consider a general finite graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ . For any  $n \in \{0, \dots, |\mathcal{V}|\}$ , we may consider the action of  $H_{\mathcal{G}}$ , restricted to  $\mathcal{H}_{\mathcal{V}}^{\text{mag}}(n)$ , in some coordinates.

From definition (1) we know  $H_{\mathcal{G}}$  is a self-adjoint operator on  $\mathcal{H}_{\mathcal{V}}$ . For any  $E_1, E_2 \in \mathbb{R}$  with  $E_1 \leq E_2$ , let  $\mathfrak{L}_{\mathcal{G}}(E_1, E_2)$  denote the subspace

$$\mathfrak{L}_{\mathcal{G}}(E_1, E_2) = \bigoplus_{\lambda \in \text{spec}(H_{\mathcal{G}}) \cap [E_1, E_2]} \ker(H_{\mathcal{G}} - \lambda \mathbb{1}). \quad (14)$$



Note that by Proposition 2.3, if  $E_1 \leq 0 \leq E_2$ , then  $\mathfrak{L}_{\mathcal{G}}(E_1, E_2) = \mathfrak{L}_{\mathcal{G}}(0, E_2)$ .

Now we state the approximate eigenvectors.

**Definition 3.5** Recalling the definition in (3), we define  $T_{\mathcal{V}}^{(n)} : \ell^2(\mathcal{V}^n) \rightarrow \mathcal{H}_{\mathcal{V}}$  by

$$T_{\mathcal{V}}^{(n)} F = \frac{1}{\sqrt{n!}} \sum_{(x_1, \dots, x_n) \in \mathcal{V}^n} F(x_1, \dots, x_n) \Omega_{\mathcal{V}}^{(n)}(x_1, \dots, x_n).$$

**Definition 3.6** For each  $\xi \in \mathbb{R}$  define the function  $f(\xi, \cdot) : \mathbb{Z} \rightarrow \mathbb{R}$  as follows: for  $\xi \in \mathbb{R} \setminus \{0\}$ , define  $f(\xi, r) = 2^{1/2} \cos(\pi \xi [r - \frac{1}{2}])$  for all  $r \in \mathbb{Z}$ ; and  $f(0, r) = 1$  for all  $r \in \mathbb{Z}$ .

We denote an element of  $\{0, 1, \dots\}^d$  as  $\kappa = (\kappa_1, \dots, \kappa_d)$ . Sometimes we need several vectors, such as  $\kappa_1, \dots, \kappa_n \in \{0, 1, \dots\}^d$ . In this case, the coordinates will be expressed as  $\kappa_k = (\kappa_{k,1}, \dots, \kappa_{k,d})$  for each  $k \in \{1, \dots, n\}$  (and similarly for similar cases).

**Definition 3.7** Fix  $d, n \in \mathbb{N}$ . For  $\kappa_1, \dots, \kappa_n \in \{0, 1, \dots\}^d$ , and each  $N \in \mathbb{N}$ , we define a function,  $F_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n; \cdot) \in \ell^2(\Lambda(d, N)^n)$ , by the formula

$$F_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n; \mathbf{r}_1, \dots, \mathbf{r}_n) = [L(d, N)]^{-nd/2} \sum_{\pi \in S_n} \prod_{k=1}^n \prod_{j=1}^d f([L^+(d, N)]^{-1} \kappa_{k,j}, r_{\pi_k, j}), \quad (15)$$

which is in the range of  $\mathfrak{S}_{\Lambda(d, N)}^{(n)}$  because of the averaging over the action of  $S_n$ . Then we define a vector,  $\tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n) \in \mathcal{H}_{\Lambda(d, N)}^{\text{mag}}(n)$ , by the formula  $\tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n) = T_{\Lambda(d, N)}^{(n)} F_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n; \cdot)$ .

We put the tilde over the vector because these are not actual eigenvectors. But we will see that they are “approximate eigenvectors.” One immediate corollary of the definition is this:

**Lemma 3.8** Fix  $d, n \in \mathbb{N}$ . For any choice of  $\kappa_1, \dots, \kappa_n \in \{0, 1, \dots\}^d$ ,

$$S_{\Lambda(d, N)}^- \tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n) = [L(d, N)]^{d/2} \tilde{\Psi}_{d,N}^{(n+1)}(\kappa_1, \dots, \kappa_n, \mathbf{0}). \quad (16)$$

**Proof:** For a general graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , we have  $T_{\mathcal{V}}^{(n+1)} \hat{S}_{\mathcal{V}}^{(n, n+1)} = S_{\mathcal{V}}^- T_{\mathcal{V}}^{(n)}$ , where  $\hat{S}_{\mathcal{V}}^{(n, n+1)} : \ell^2(\mathcal{V}^n) \rightarrow \ell^2(\mathcal{V}^{n+1})$  is the linear transformation such that

$$\hat{S}_{\mathcal{V}}^{(n, n+1)} F(x_1, \dots, x_{n+1}) = F(x_1, \dots, x_n) + F(x_2, \dots, x_{n+1}) + \sum_{k=2}^n F(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_{n+1}).$$

But then, from (15), we see that this is equivalent to (16).  $\square$

For any  $\pi \in S_n$ , we have

$$\tilde{\Psi}_{d,N}^{(n)}(\kappa_{\pi_1}, \dots, \kappa_{\pi_n}) = \tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n), \quad (17)$$

which we see directly by inspection of Definition 3.5 and Definition 3.7. For  $d, n \in \mathbb{N}$ , let

$$\mathcal{K}(d, n) \stackrel{\text{def}}{=} (\{0, 1, \dots\}^d)^n = \{(\kappa_1, \dots, \kappa_n) : \kappa_1, \dots, \kappa_n \in \{0, 1, \dots\}^d\}.$$

Given any  $(\kappa_1, \dots, \kappa_n) \in \mathcal{K}(d, n)$ , define

$$[(\kappa_1, \dots, \kappa_n)]_{S_n} = \{(\kappa_{\pi_1}, \dots, \kappa_{\pi_n}) : \pi \in S_n\} \subset \mathcal{K}(d, n).$$

Let us define

$$\tilde{\mathcal{K}}(d, n) = \{[(\kappa_1, \dots, \kappa_n)]_{S_n} : (\kappa_1, \dots, \kappa_n) \in \mathcal{K}(d, n)\}, \quad (18)$$

which is the quotient space  $\mathcal{K}(d, n)/\sim$ , where the equivalence relation is:  $(\kappa_1, \dots, \kappa_n) \sim (\kappa'_1, \dots, \kappa'_n)$  if and only if  $[(\kappa_1, \dots, \kappa_n)]_{S_n} = [(\kappa'_1, \dots, \kappa'_n)]_{S_n}$ .

An easy variational calculation follows. Let us define a quantity (related to the spectral gap)

$$\gamma \stackrel{\text{def}}{=} \pi^2/2. \quad (19)$$

Then we have the following:

**Proposition 3.9** *Fix  $d, n \in \mathbb{N}$ . For any choice of  $\kappa_1, \dots, \kappa_n \in \{0, 1, \dots\}^d$ ,*

$$\lim_{N \rightarrow \infty} \left\| \left( \gamma^{-1} [L(d, N)]^2 H_{\Lambda(d, N)} - \sum_{k=1}^n \sum_{j=1}^d \kappa_{k,j}^2 \right) \tilde{\Psi}_{d, N}^{(n)}(\kappa_1, \dots, \kappa_n) \right\| = 0. \quad (20)$$

Moreover, for any additional choice of  $\kappa'_1, \dots, \kappa'_n \in \{0, 1, \dots\}^d$ ,

$$\lim_{N \rightarrow \infty} \langle \tilde{\Psi}_{d, N}^{(n)}(\kappa_1, \dots, \kappa_n), \tilde{\Psi}_{d, N}^{(n)}(\kappa'_1, \dots, \kappa'_n) \rangle = n! \sum_{\pi \in S_n} \prod_{k=1}^n \mathbf{1}_{\{\kappa_{\pi_k} = \kappa'_k\}}. \quad (21)$$

We will prove this in Appendix D. It is easy, using the explicit formulas for the eigenvectors.

From (17) and (21), we see that the appropriate labeling of approximate eigenstates is given by elements of  $\tilde{\mathcal{K}}(d, n)$  from (18). But from (20), we also see that the appropriate energy for  $\tilde{\Psi}_{d, N}^{(n)}(\kappa_1, \dots, \kappa_n)$ , modulo the scale  $\gamma \cdot [L(d, N)]^{-2}$ , is  $\sum_{k=1}^n \sum_{j=1}^d \kappa_{k,j}^2$  for  $(\kappa_1, \dots, \kappa_n) \in \mathcal{K}(d, n)$ . For each  $m \in \{0, 1, \dots\}$ , let us define

$$\mathcal{K}(d, n, m) \stackrel{\text{def}}{=} \left\{ (\kappa_1, \dots, \kappa_n) \in \mathcal{K}(d, n) : \sum_{k=1}^n \sum_{j=1}^d \kappa_{k,j}^2 = m \right\}. \quad (22)$$

Let us also define, for each  $m \in \{0, 1, \dots\}$ ,

$$\tilde{\mathcal{K}}(d, n, m) = \{[(\kappa_1, \dots, \kappa_n)]_{S_n} : (\kappa_1, \dots, \kappa_n) \in \mathcal{K}(d, n, m)\}. \quad (23)$$

Finally, let us define the number  $R(d, n, m) \in \{0, 1, \dots\}$ ,

$$R(d, n, m) = |\tilde{\mathcal{K}}(d, n, m)|, \quad (24)$$

which is supposed to enumerate the linearly independent (indeed orthogonal) eigenvectors whose eigenvalues, modulo the scale  $\gamma \cdot [L(d, N)]^{-2}$ , are all near to  $m$ . Now we state the main theorem for the linear spin wave approximation that we will use.

**Theorem 3.10** *Fix  $d, n \in \mathbb{N}$ . For each  $N \in \mathbb{N}$ , we may define vectors  $\Psi_{d, N}^{(n)}(m, r)$ , for each  $m \in \{0, 1, \dots\}$  and each  $r \in \{1, \dots, R(d, n, m)\}$ , which are all orthogonal, and such that, for each  $m \in \{0, 1, \dots\}$  and each  $\epsilon \in (0, 1/2)$ , there is an integer  $N_1(d, n, m, \epsilon) \in \mathbb{N}$  such that for all  $N \geq N_1(d, n, m, \epsilon)$  the following properties hold.*

- (i) *For each  $m_1 \in \{0, \dots, m\}$  and each  $r \in \{1, \dots, R(d, n, m_1)\}$ , the vector  $\Psi_{d, N}^{(n)}(m_1, r)$  is an eigenvector of  $\gamma^{-1} [L(d, N)]^2 H_{\Lambda(d, N)}$  with associated eigenvalue  $E_{d, N}^{(n)}(m_1, r) \in (m_1 - \epsilon, m_1 + \epsilon)$ .*
- (ii) *The vectors  $\{\Psi_{d, N}^{(n)}(m_1, r) : m_1 \in \{0, \dots, m\}, r \in \{1, \dots, R(d, n, m_1)\}\}$  are an orthonormal basis for  $\mathfrak{L}_{\Lambda(d, N)}(0, \gamma^{-1} [L(d, N)]^2 (m + \epsilon)) \cap \mathcal{H}_{\Lambda(d, N)}^{\text{mag}}(n)$ , using the definition from (14).*

## 4 Conditional Proof of the Main Theorem

The following easy result is the first step to a calculation technique.

**Lemma 4.1** *For any finite graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  and any  $n \in \{0, \dots, \lfloor \frac{1}{2}|\mathcal{V}| \rfloor\}$ :  
(1) for each choice of  $E_1, E_2 \in \mathbb{R}$  with  $0 \leq E_1 \leq E_2$ , we have*

$$\dim(\mathcal{L}_{\mathcal{G}}(E_1, E_2) \cap \mathcal{H}_{\mathcal{V}}^{\text{mag}}(n)) \geq \dim(\mathcal{L}_{\mathcal{G}}(E_1, E_2) \cap \mathcal{H}_{\mathcal{V}}^{\text{mag}}(n-1));$$

(2) moreover, we may calculate  $\mathfrak{E}_n(\mathcal{G})$  from (11) as

$$\mathfrak{E}_n(\mathcal{G}) = \min(\{E \in [0, \infty) : \dim(\mathcal{L}_{\mathcal{G}}(0, E) \cap \mathcal{H}_{\mathcal{V}}^{\text{mag}}(n)) > \dim(\mathcal{L}_{\mathcal{G}}(0, E) \cap \mathcal{H}_{\mathcal{V}}^{\text{mag}}(n-1))\}).$$

**Proof:** By Lemma 2.2, we know that  $S_{\mathcal{V}}^-$  is an isomorphism of  $\mathcal{H}_{\mathcal{V}}(m, n)$  onto  $\mathcal{H}_{\mathcal{V}}(m+1, n)$  for each  $m \in \{n, \dots, |\mathcal{V}| - n - 1\}$ . Since this operator commutes with  $H_{\mathcal{G}}$ , we may simplify the definition of  $\mathfrak{E}_n(\mathcal{G})$  from (11):

$$\mathfrak{E}_n(\mathcal{G}) = \min \text{spec} (H_{\mathcal{G}} \upharpoonright \mathcal{H}_{\mathcal{V}}(n, n)).$$

But, then we see that, again by Lemma 2.2,

$$\mathcal{H}_{\mathcal{V}}(n, n) = \mathcal{H}_{\mathcal{V}}^{\text{mag}}(n) \cap \left( S_{\mathcal{V}}^-(\mathcal{H}_{\mathcal{V}}^{\text{mag}}(n-1)) \right)^\perp.$$

Also, by Lemma 2.2, we know that  $S_{\mathcal{V}}^-$  is an isomorphism of  $\mathcal{H}_{\mathcal{V}}^{\text{mag}}(n-1)$  onto  $S_{\mathcal{V}}^-(\mathcal{H}_{\mathcal{V}}^{\text{mag}}(n-1))$ . So, again, since  $S_{\mathcal{V}}^-$  commutes with  $H_{\mathcal{G}}$ , we may conclude both facts stated in the lemma.  $\square$

**Corollary 4.2** *Fix  $d, n \in \{1, 2, \dots\}$ . Define  $\mathfrak{m}(d, n) \in \{1, 2, \dots\}$  as*

$$\mathfrak{m}(d, n) \stackrel{\text{def}}{=} \min(\{m \in \{1, 2, \dots\} : R(d, n, m) > R(d, n-1, m)\}).$$

*For any  $\epsilon > 0$ , and  $N \geq N_1(d, n, \mathfrak{m}(d, n), \epsilon)$  (as in Theorem 3.10), we have*

$$|\gamma^{-1}[L(d, N)]^2 \mathfrak{E}_n(\Lambda(d, N)) - \mathfrak{m}(d, n)| \leq \epsilon.$$

Before proving this simple corollary of Theorem 3.10 and Lemma 4.1, let us note this formula:

**Lemma 4.3** *For each  $d, n \in \{1, 2, \dots\}$ , we have*

$$\mathfrak{m}(d, n) = n. \tag{25}$$

**Proof:** For any  $m < n$ , choose an element of  $\tilde{\mathcal{K}}(d, n, m)$ . We may find some  $(\kappa_1, \dots, \kappa_n) \in \mathcal{K}(d, n, m)$  such that the chosen element is  $[(\kappa_1, \dots, \kappa_n)]_{S_n}$ . But then  $\|\kappa_1\|^2 + \dots + \|\kappa_n\|^2 = m$ . The minimum nonzero value for  $\|\kappa_j\|^2$  is 1 for each  $j$ . Thus,  $m < n$ . So  $\kappa_j = \mathbf{0}$  for some  $j \in \{1, \dots, n\}$ . Without loss of generality (due to the permutation symmetry), we may assume that the  $(\kappa_1, \dots, \kappa_n)$  that we found had  $\kappa_n = \mathbf{0}$ . This means that  $[(\kappa_1, \dots, \kappa_{n-1})]_{S_{n-1}}$  is an element of  $\tilde{\mathcal{K}}(d, n-1, m)$ . Since we have merely removed one zero element, and we mod out by the action of the symmetric group in the quotient space, this mapping is a bijection. So  $R(d, n-1, m) = R(d, n, m)$  for each  $m < n$ . More generally, even for  $m \geq n$ , there is always a bijection between the set of elements of  $\tilde{\mathcal{K}}(d, n, m)$  having at least one  $\mathbf{0}$  part and the set of elements of  $\tilde{\mathcal{K}}(d, n-1, m)$ . But, on the other hand, for  $\delta_{d,1} \stackrel{\text{def}}{=} (1, 0, \dots, 0) \in \{0, 1, \dots\}^d$ , we may see that  $[(\delta_{d,1}, \dots, \delta_{d,1})]_{S_n}$  is an element of  $\tilde{\mathcal{K}}(d, n, n)$  not equal to  $[(\kappa_1, \dots, \kappa_{n-1}, \mathbf{0})]_{S_n}$  for any choice of  $[(\kappa_1, \dots, \kappa_{n-1})]_{S_{n-1}}$  in  $\tilde{\mathcal{K}}(d, n-1, m)$ . So  $\tilde{\mathcal{K}}(d, n, n)$  is not bijective to  $\tilde{\mathcal{K}}(d, n-1, m)$ . Rather,  $R(d, n, n) > R(d, n-1, n)$ .  $\square$

**Proof of Corollary 4.2:** Let  $\widehat{\mathfrak{L}}_{\Lambda(d,N)}(E_1, E_2) = \mathfrak{L}_{\Lambda(d,N)}(\gamma \cdot [L(d, N)]^{-2} E_1, \gamma \cdot [L(d, N)]^{-2} E_2)$ . Let  $m = \mathfrak{m}(d, n)$  (which is  $n$ ). We know, by Theorem 3.10, that for any  $m_1 \in \{1, 2, \dots\}$  with  $m_1 < m$ , we have

$$\dim \left( \widehat{\mathfrak{L}}_{\Lambda(d,N)}(0, m_1 + \epsilon) \cap \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n) \right) = \sum_{m_2=0}^{m_1} R(d, n, m_2).$$

But since  $m_1 < m$ , we have that this is also equal to  $\sum_{m_2=0}^{m_1} R(d, n-1, m_2)$ . Therefore, we have

$$\dim \left( \widehat{\mathfrak{L}}_{\Lambda(d,N)}(0, m_1 + \epsilon) \cap \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n) \right) = \dim \left( \widehat{\mathfrak{L}}_{\Lambda(d,N)}(0, m_1 + \epsilon) \cap \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n-1) \right).$$

So  $\gamma^{-1} \cdot [L(d, N)]^2 \mathfrak{E}_n(\Lambda(d, N)) > m_1 + \epsilon$ , by Lemma 4.1. By Theorem 3.10, we also know that  $\dim(\widehat{\mathfrak{L}}_{\Lambda(d,N)}(m-1+\epsilon, m-\epsilon) \cap \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n))$  equals 0. Hence,  $\gamma^{-1} \cdot [L(d, N)]^2 \mathfrak{E}_n(\Lambda(d, N)) \geq m - \epsilon$ . But we also know, by Theorem 3.10,

$$\begin{aligned} \dim \left( \widehat{\mathfrak{L}}_{\Lambda(d,N)}(0, m + \epsilon) \cap \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n) \right) &= \sum_{m_1=0}^{m-1} R(d, n, m_1) + R(d, n, m) \\ &= \sum_{m_1=0}^{m-1} R(d, n-1, m_1) + R(d, n, m) \\ &> \sum_{m_1=0}^{m-1} R(d, n, m_1) + R(d, n-1, m) \\ &= \dim \left( \widehat{\mathfrak{L}}_{\Lambda(d,N)}(0, m + \epsilon) \cap \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n-1) \right). \end{aligned}$$

So  $\gamma^{-1}[L(d, N)]^2 \mathfrak{E}_n(\Lambda(d, N)) \leq m + \epsilon$ . □

Combining Corollary 4.2 and Lemma 4.3, we conclude the following.

**Corollary 4.4** Fix  $d, n \in \{1, 2, \dots\}$ . Then,

$$\mathfrak{E}_n(\Lambda(d, N)) \sim \gamma \cdot [L(d, N)]^{-2} \cdot n, \quad \text{as } N \rightarrow \infty.$$

**Proof:** Corollary 4.2 and equation (25) from Lemma 4.3 imply that, for each  $\epsilon > 0$ , we have  $\mathfrak{E}_n(\Lambda(d, N))/(\gamma \cdot [L(d, N)]^{-2} \cdot n)$  is in the interval  $[1 - \epsilon, 1 + \epsilon]$  for sufficiently large  $N$ : specifically,  $N > N_1(d, n, \epsilon)$ , from Theorem 3.10. That is the definition of asymptotic equivalence. □

Recall from Definition 3.4 that  $L(d, N) = \lfloor N^{1/d} \rfloor$ . In particular, this means  $L(d, N) \sim N^{1/d}$ , as  $N \rightarrow \infty$ . So, Corollary 4.4 may be rewritten as

$$\mathfrak{E}_n(\Lambda(d, N)) \sim n \gamma N^{-2/d}, \quad \text{as } N \rightarrow \infty. \quad (26)$$

We will use (26). We also need the following:

**Lemma 4.5** Suppose that  $(t_a, t_{a+1}, \dots)$  is a sequence of strictly positive numbers such that  $t_N \sim CN^{-p}$  for some  $C \in (0, \infty)$  and some  $p > 0$ . Let us define a function  $\nu : \{a, a+1, \dots\} \rightarrow \{a, a+1, \dots\} \cup \{\infty\}$  by the formula

$$\nu(N) = \inf\{N' \in \{N, N+1, \dots\} : t_{N'} = \min\{t_a, \dots, t_{N'}\}\}.$$

Then  $t_{\nu(N)} \sim CN^{-p}$ .

This is an exercise in advanced calculus. For completeness we include its proof in Appendix A. Now we may give the conditional proof of our main theorem.

**Proof of Theorem 2.6:** This follows from Corollary 4.4, as rewritten in equation (26), as well as Lemma 4.5, and Proposition 3.3. For each  $n$ , let  $\nu_n$  be defined as

$$\nu_n(N) = \inf\{N' \in \{N, N+1, \dots\} : \mathfrak{E}_n(\Lambda(d, N')) = \min\{\mathfrak{E}_n(\Lambda(d, 2n)), \dots, \mathfrak{E}_n(\Lambda(d, N'))\}\}. \quad (27)$$

Then we know that  $\mathfrak{E}_n(\Lambda(d, \nu_n(N))) \sim C_n N^{-p}$  by (26), where  $C_n = n\gamma$  and  $p = 2/d$ . The hypotheses of Proposition 3.3 are satisfied for  $n$  and the finite sequence of graphs

$$(\Lambda(d, 2n), \dots, \Lambda(d, \nu_n(N)))$$

by (27). So we know that

$$\min_{N' \in \{2n, \dots, \nu_n(N)\}} \min_{n' \in \{n, \dots, \lfloor N'/2 \rfloor\}} \mathfrak{E}_{n'}(\Lambda(d, N')) = \mathfrak{E}_n(\Lambda(d, \nu_n(N))).$$

In particular, since  $\nu_n(N) \geq N$ , we have

$$\min_{n' \in \{n, \dots, \lfloor N/2 \rfloor\}} \mathfrak{E}_{n'}(\Lambda(d, N)) \geq \mathfrak{E}_n(\Lambda(d, \nu_n(N))).$$

But as we already established,  $\mathfrak{E}_n(\Lambda(d, \nu_n(N))) \sim C_n N^{-p}$ . Therefore, defining a new quantity  $\mathfrak{E}_{n\uparrow}^{\min}(\Lambda(d, N))$  as

$$\mathfrak{E}_{n\uparrow}^{\min}(\Lambda(d, N)) = \min_{n' \in \{n, \dots, \lfloor N/2 \rfloor\}} \mathfrak{E}_{n'}(\Lambda(d, N)), \quad (28)$$

we actually have  $\liminf_{N \rightarrow \infty} \mathfrak{E}_{n\uparrow}^{\min}(\Lambda(d, N)) / (C_n N^{-p}) \geq 1$ . But also,  $\mathfrak{E}_{n\uparrow}^{\min}(\Lambda(d, N)) \leq \mathfrak{E}_n(\Lambda(d, N))$  because  $\mathfrak{E}_n(\Lambda(d, N))$  is one of the terms in the minimum on the right hand side (28). So since  $\mathfrak{E}_n(\Lambda(d, N)) \sim C_n N^{-p}$ , as  $N \rightarrow \infty$ , we also have

$$\limsup_{N \rightarrow \infty} \frac{\mathfrak{E}_{n\uparrow}^{\min}(\Lambda(d, N))}{C_n N^{-p}} \leq 1.$$

In other words, putting these two inequalities together,

$$\mathfrak{E}_{n\uparrow}^{\min}(\Lambda(d, N)) \sim C_n N^{-p}, \text{ as } N \rightarrow \infty.$$

But now this allows us to get the result we want, because this is true for all  $n$ . In particular, fixing  $n$ , it is also true for  $n+1$ . So

$$\mathfrak{E}_{n+1\uparrow}^{\min}(\Lambda(d, N)) \sim C_{n+1} N^{-p}.$$

But  $C_{n+1} = (n+1)\gamma$  is strictly greater than  $C_n = n\gamma$ . From the definition (28) again,

$$\mathfrak{E}_{n\uparrow}^{\min}(\Lambda(d, N)) = \min\{\mathfrak{E}_n(\Lambda(d, N)), \mathfrak{E}_{n+1\uparrow}^{\min}(\Lambda(d, N))\}.$$

But  $\mathfrak{E}_n(\Lambda(d, N)) \sim C_n N^{-p}$  and  $\mathfrak{E}_{n+1\uparrow}^{\min}(\Lambda(d, N)) \sim C_{n+1} N^{-p}$ . The strict inequality  $C_n < C_{n+1}$  implies that the former is strictly smaller than the latter for sufficiently large  $N$ . So

$$\mathfrak{E}_{n\uparrow}^{\min}(\Lambda(d, N)) = \mathfrak{E}_n(\Lambda(d, N)), \text{ and } \mathfrak{E}_{n\uparrow}^{\min}(\Lambda(d, N)) < \mathfrak{E}_{n+1\uparrow}^{\min}(\Lambda(d, N)),$$

for sufficiently large  $N$ . Decoding the definition (28) again, this gives the desired result.  $\square$

## 5 Proof of Key Step I: Inductive argument for FOEL

In this section we will prove Proposition 3.1 and Proposition 3.3. For Proposition 3.1 we will actually prove a generalization, where the generalization is that we do not consider just the Heisenberg model on general graphs with all coupling constants equal to 1. But we allow the coupling constants, themselves, to be variable.

**Definition 5.1** *Given a graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , let us say that  $J$  is a “valid choice of coupling coefficients for  $\mathcal{G}$ ” if  $J$  denotes a function  $J : \mathcal{E} \rightarrow \mathbb{R}$ , such that  $J(\{x, y\})$  is nonnegative for each  $\{x, y\} \in \mathcal{E}$ . We sometimes say “ $J$  is valid for  $\mathcal{G}$ .”*

Then we define the Heisenberg Hamiltonian  $H(\mathcal{G}, J)$  on  $\mathcal{H}_{\mathcal{V}}$  as

$$H(\mathcal{G}, J) = \sum_{\{x, y\} \in \mathcal{E}} J(\{x, y\}) h_{xy}, \quad (29)$$

where  $h_{xy}$  is just as in the original definition (1). We define  $\mathfrak{E}_n(\mathcal{G}, J)$  as

$$\mathfrak{E}_n(\mathcal{G}, J) = \min \text{spec} \left( H(\mathcal{G}, J) \upharpoonright \mathcal{H}_{\mathcal{V}}^{\text{spin}}(n) \right), \quad (30)$$

analogously to the definition in (11).

**Definition 5.2** *We say that  $(\mathcal{G}, J)$  satisfies FOEL- $n$  if  $\mathfrak{E}_n(\mathcal{G}, J) \leq \mathfrak{E}_r(\mathcal{G}, J)$  for all  $r \geq n$ , i.e., for all  $r \in \{n, \dots, \lfloor \frac{1}{2}|\mathcal{V}| \rfloor\}$ .*

We note that Lemma 2.2 remains true if we replace  $H(\mathcal{G})$  by  $H(\mathcal{G}, J)$  for a valid choice of couplings. The analogue of Proposition 2.3 is as follows. Given  $(\mathcal{G}, J)$  with  $J$  valid for  $\mathcal{G}$ , define a new graph  $\mathcal{G}_J = (\mathcal{V}, \mathcal{E}_J)$  with  $\mathcal{E}_J = \{\{x, y\} \in \mathcal{E} : J(\{x, y\}) > 0\}$ . Then  $(\mathcal{G}, J)$  satisfies strict FOEL-0 if  $\mathcal{G}_J$  is connected. Lieb and Mattis prove all of their results in [20] including this one using general couplings. We merely stated Proposition 2.3 for the special case that all couplings are equal to 1 for ease of exposition in the introduction. In particular, the analogue of equation (7) is true: if  $J$  is a valid choice of coupling coefficients for  $\mathcal{G}$ , then

$$H(\mathcal{G}, J) \geq 0. \quad (31)$$

The following lemma is a key to the inductive proof of FOEL. In order to simplify notation, let us begin to use the following natural convention:

**Definition 5.3** *If  $n \in \{1, 2, \dots\}$  satisfies  $n > \lfloor \frac{1}{2}|\mathcal{V}| \rfloor$ , then let us define  $\mathfrak{E}_n(\mathcal{G}, J)$  to be  $+\infty$ , interpreting the empty minimum in (30) as  $+\infty$ . Similarly, for the case of constant coupling coefficients 1, we also define  $\mathfrak{E}_n(\mathcal{G})$  to be  $+\infty$ , in this case that  $n$  is larger than the maximum possible value to get  $\mathcal{H}_{\mathcal{V}}^{\text{spin}}(n)$  different than  $\{0\}$ .*

**Lemma 5.4** *Suppose that  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  is a finite graph, and that  $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$  is a graph such that  $\mathcal{V}' = \mathcal{V} \cup \{x'\}$  for a single vertex  $x'$  not in  $\mathcal{V}$ , and such that  $\mathcal{E} \subseteq \mathcal{E}'$ . Let  $J$  be a valid choice of coupling coefficients for  $\mathcal{G}$ , and let  $J'$  be a valid choice of coupling coefficients for  $\mathcal{G}'$ . Suppose, moreover, that  $J'(\{x, y\}) \geq J(\{x, y\})$  for all edges  $\{x, y\} \in \mathcal{E}$ . Then  $\mathfrak{E}_n(\mathcal{G}', J') \geq \min\{\mathfrak{E}_n(\mathcal{G}, J), \mathfrak{E}_{n-1}(\mathcal{G}, J)\}$  for each  $n \in \{1, \dots, \lfloor \frac{1}{2}|\mathcal{V}'| \rfloor\}$ .*

**Proof:** Note that there is a canonical isomorphism  $\mathcal{H}_{\mathcal{V}'} \cong \mathcal{H}_{\mathcal{V}} \otimes \mathbb{C}^2$ , because the last spin site  $x'$  also has spin Hilbert space equivalent to  $\mathbb{C}^2$ . By addition of angular momenta, using the definition in Lemma 2.2, and the formula for addition of angular momenta in SU(2)

$$\mathcal{H}_{\mathcal{V}'}^{\text{spin}}(n) \subseteq (\mathcal{H}_{\mathcal{V}}^{\text{spin}}(n) \oplus \mathcal{H}_{\mathcal{V}}^{\text{spin}}(n-1)) \otimes \mathbb{C}^2, \quad (32)$$

for each  $n \in \{0, \dots, \lfloor s_{\max}(\mathcal{V}') \rfloor\}$ , and where we define  $\mathcal{H}_{\mathcal{V}}^{\text{spin}}(-1) = \{0\}$  for consistency, and we define  $\mathcal{H}_{\mathcal{V}}^{\text{spin}}(n) = \{0\}$  if  $n > \lfloor \frac{1}{2}|\mathcal{V}| \rfloor$ . See for example [16] for a proof of “addition of angular momenta” for  $\text{SU}(2)$  representations.

Now we may choose a vector  $\psi \in \mathcal{H}_{\mathcal{V}'}^{\text{spin}}(n)$  with  $\|\psi\| = 1$  and such that  $H(\mathcal{G}', J')\psi = E_n(\mathcal{G}', J')\psi$  because  $E_n(\mathcal{G}', J')$  is the minimum of the spectrum of  $H(\mathcal{G}', J')$  restricted to the subspace  $\mathcal{H}_{\mathcal{V}'}^{\text{spin}}(n)$ . Moreover, by (32), we may find two orthogonal vectors  $\psi_1 \in \mathcal{H}_{\mathcal{V}}^{\text{spin}}(n) \otimes \mathbb{C}^2$  and  $\psi_2 \in \mathcal{H}_{\mathcal{V}}^{\text{spin}}(n-1) \otimes \mathbb{C}^2$  such that  $\psi = \psi_1 + \psi_2$ . Furthermore, since  $J'(\{x, y\}) \geq J(\{x, y\})$  for all  $\{x, y\} \in \mathcal{E}$  (and since  $J'(\{x, y\}) \geq 0$  for all  $\{x, y\} \in \mathcal{E}' \setminus \mathcal{E}$ ), we have, in part using equation (31),

$$H(\mathcal{G}', J') \geq H(\mathcal{G}, J) \otimes \mathbb{1}_{\mathbb{C}^2},$$

again, viewing  $\mathcal{H}_{\mathcal{V}'}$  as equivalent to  $\mathcal{H}_{\mathcal{V}} \otimes \mathbb{C}^2$ . Therefore, we have

$$\mathfrak{E}_n(\mathcal{G}', J') = \langle \psi, H(\mathcal{G}', J')\psi \rangle \geq \langle \psi, (H(\mathcal{G}, J) \otimes \mathbb{1}_{\mathbb{C}^2})\psi \rangle. \quad (33)$$

But both of the subspaces,  $\mathcal{H}_{\mathcal{V}}^{\text{spin}}(n) \otimes \mathbb{C}^2$  and  $\mathcal{H}_{\mathcal{V}}^{\text{spin}}(n-1) \otimes \mathbb{C}^2$ , are invariant subspaces for  $H(\mathcal{G}, J) \otimes \mathbb{1}_{\mathbb{C}^2}$ . Hence, we may actually write

$$\langle \psi, (H(\mathcal{G}, J) \otimes \mathbb{1}_{\mathbb{C}^2})\psi \rangle = \langle \psi_1, (H(\mathcal{G}, J) \otimes \mathbb{1}_{\mathbb{C}^2})\psi_1 \rangle + \langle \psi_2, (H(\mathcal{G}, J) \otimes \mathbb{1}_{\mathbb{C}^2})\psi_2 \rangle. \quad (34)$$

But by the definition of  $\mathfrak{E}_n(\mathcal{G}, J)$  as the minimum of the spectrum of  $H(\mathcal{G}, J)$  restricted to the subspace  $\mathcal{H}_{\mathcal{V}}^{\text{spin}}(n)$  and similarly for  $\mathfrak{E}_{n-1}(\mathcal{G}, J)$ , we have

$$\langle \psi_1, (H(\mathcal{G}, J) \otimes \mathbb{1}_{\mathbb{C}^2})\psi_1 \rangle \geq \mathfrak{E}_n(\mathcal{G}, J)\|\psi_1\|^2, \quad \langle \psi_2, (H(\mathcal{G}, J) \otimes \mathbb{1}_{\mathbb{C}^2})\psi_2 \rangle \geq \mathfrak{E}_{n-1}(\mathcal{G}, J)\|\psi_2\|^2.$$

Therefore, combining this with (34), we have

$$\langle \psi, (H(\mathcal{G}, J) \otimes \mathbb{1}_{\mathbb{C}^2})\psi \rangle \geq \min\{\mathfrak{E}_n(\mathcal{G}, J), \mathfrak{E}_{n-1}(\mathcal{G}, J)\}(\|\psi_1\|^2 + \|\psi_2\|^2). \quad (35)$$

By orthogonality,  $\|\psi_1\|^2 + \|\psi_2\|^2 = \|\psi\|^2 = 1$ . So combining (33) with (35) gives the result.  $\square$

**Proposition 5.5** *Suppose that  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  is a finite graph, and that  $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$  is a graph such that  $\mathcal{V}' = \mathcal{V} \cup \{x'\}$  for a single vertex  $x'$  not in  $\mathcal{V}$ , and such that  $\mathcal{E} \subseteq \mathcal{E}'$ . Let  $J$  be valid for  $\mathcal{G}$ , and let  $J'$  be valid for  $\mathcal{G}'$ .*

*Suppose for some  $n \in \{0, 1, \dots, \lfloor \frac{1}{2}|\mathcal{V}| \rfloor\}$ , that  $(\mathcal{G}, J)$  satisfies FOEL- $n$ . Suppose, further that the following two conditions are satisfied:*

- (1)  $J'(\{x, y\}) \geq J(\{x, y\})$  for all edges  $\{x, y\} \in \mathcal{E}$ .
- (2)  $\mathfrak{E}_n(\mathcal{G}', J') \leq \mathfrak{E}_n(\mathcal{G}, J)$ .

*Then  $(\mathcal{G}', J')$  satisfies FOEL- $n$ .*

**Proof:** Suppose  $r \geq n+1$  is fixed. Applying Lemma 5.4 (with  $n$  in that lemma replaced by  $r$ ),

$$\mathfrak{E}_r(\mathcal{G}', J') \geq \min\{\mathfrak{E}_r(\mathcal{G}, J), \mathfrak{E}_{r-1}(\mathcal{G}, J)\}.$$

But by the assumption that  $(\mathcal{G}, J)$  satisfies FOEL- $n$ , we know that

$$\min_{r \geq n+1} \min\{\mathfrak{E}_r(\mathcal{G}, J), \mathfrak{E}_{r-1}(\mathcal{G}, J)\} \geq \mathfrak{E}_n(\mathcal{G}, J).$$

Therefore, we conclude

$$\min_{r \geq n+1} \mathfrak{E}_r(\mathcal{G}', J') \geq \min_{r \geq n+1} \min\{\mathfrak{E}_r(\mathcal{G}, J), \mathfrak{E}_{r-1}(\mathcal{G}, J)\} \geq \mathfrak{E}_n(\mathcal{G}, J) \geq \mathfrak{E}_n(\mathcal{G}', J').$$

But this is a restatement of FOEL- $n$  from Definition 5.2.  $\square$

Obviously, the proof of Proposition 3.1 follows.

**Proof of Proposition 3.1:** Use Proposition 5.5, with  $J(\{x, y\}) = 1$  for all  $\{x, y\} \in \mathcal{E}$  and  $J'(\{x, y\}) = 1$  for all  $\{x, y\} \in \mathcal{E}'$ . It is easy to see that the hypotheses of Proposition 5.5 are satisfied for these choices, given the hypotheses of Proposition 3.1.  $\square$

We can also, immediately prove the following corollary

**Corollary 5.6** *For any  $n \in \mathbb{N}$  and for any  $N \in \{n+1, n+2, \dots\}$ , suppose that  $\mathcal{G}_{2n}, \mathcal{G}_{2n+1}, \dots, \mathcal{G}_N$  forms a family of graphs, and for each  $k \in \{2n, \dots, N\}$  that  $J_k : \mathcal{E}_k \rightarrow \mathbb{R}$  is a valid choice of couplings on  $\mathcal{G}_k$ , satisfying the following conditions.*

(i) *For each  $k \in \{2n, \dots, N\}$ , we have that  $\mathcal{G}_k = (\mathcal{V}_k, \mathcal{E}_k)$  satisfies  $|\mathcal{V}_k| = k$ .*

(ii) *For each  $k \in \{2n, \dots, N-1\}$ , we have that*

(a)  $\mathcal{V}_k \subset \mathcal{V}_{k+1}$ ,

(b)  $\mathcal{E}_k \subset \mathcal{E}_{k+1}$ ,

(c)  $|\mathcal{V}_{k+1} \setminus \mathcal{V}_k| = 1$ , and

(d)  $J_{k+1}(\{x, y\}) \geq J_k(\{x, y\})$  for each  $\{x, y\} \in \mathcal{E}_k$ .

(iii) *For each  $k \in \{2n, \dots, N-1\}$ , we have that  $\mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}) \leq \mathfrak{E}_n(\mathcal{G}_k, J_k)$ .*

*Then  $(\mathcal{G}_k, J_k)$  satisfies FOEL- $n$  according to Definition 5.2 for each  $k \in \{2n, \dots, N\}$ .*

**Proof:** The proof is just like the proof of Corollary 3.2, except that now we use Proposition 5.5 everywhere that the old proof used Proposition 3.1.  $\square$

Next we will prove Proposition 3.3. It follows from Corollary 5.6.

**Definition 5.7** *Suppose that  $\mathcal{G}_n, \dots, \mathcal{G}_N$  is an ordered chain of graphs, then we say that a valid choice of coupling constants  $J_k : \mathcal{E}_k \rightarrow [0, \infty)$  for each  $k = n, \dots, N$  is a “diluted system” if*

(i)  $J_k \leq \mathbf{1}_{\mathcal{E}_k}$ , pointwise, for each  $k \in \{n, \dots, N-1\}$ , and  $J_N = \mathbf{1}_{\mathcal{E}_N}$ ;

(ii)  $J_n \leq \dots \leq J_N$  in the sense of Proposition 5.5 condition (1).

**Lemma 5.8** *If  $\mathcal{G}_{2n}, \dots, \mathcal{G}_N$  is an ordered chain of graphs, and if  $N$  is a new low of the sequence  $(\mathfrak{E}_n(\Lambda_k))_{k=2n}^N$ , then there is a “diluted system”  $J_{2n} \leq \dots \leq J_N$  such that  $(\mathfrak{E}_n(\mathcal{G}_k, J_k))_{k=2n}^N$  is monotone non-increasing, and  $J_N = \mathbf{1}_{\mathcal{E}_N}$ . In other words,  $H(\mathcal{G}_N, J_N) = H(\mathcal{G}_N)$ .*

**Proof of Lemma 5.8:** We will actually construct a diluted system  $J_{2n}, \dots, J_N$  satisfying

$$\mathfrak{E}_n(\mathcal{G}_k, J_k) = \min\{\mathfrak{E}_n(\mathcal{G}_i) : i = 2n, \dots, k\}, \quad (36)$$

for each  $k = 2n, \dots, N$ .

The proof is by induction. For the initial step, we let  $J_{2n} = \mathbf{1}_{\mathcal{E}_{2n}}$ . Then (36) follows, trivially.



Suppose that some diluted system  $J_{2n}, \dots, J_k$  has been constructed up to  $k$ , for some  $k \in \{2n, \dots, N-1\}$ , satisfying (36). For each  $t \in [0, 1]$ , we define a choice of valid coupling constants  $J_{k+1}^{(t)} : \mathcal{E}_{k+1} \rightarrow [0, \infty)$  as follows:

$$J_{k+1}^{(t)}(\{x, y\}) = \begin{cases} (1-t)J_k(\{x, y\}) + t & \text{for } \{x, y\} \in \mathcal{E}_k, \\ t & \text{for } \{x, y\} \in \mathcal{E}_{k+1} \setminus \mathcal{E}_k, \\ 0 & \text{for } \{x, y\} \notin \mathcal{E}_{k+1}. \end{cases}$$

Note that  $J_{k+1}^{(0)}$  is  $J_k$  viewed as a function on  $\mathcal{E}_{k+1}$ , as one can see from the formula. The couplings are pointwise non-decreasing in  $t$ . Also,  $J_{k+1}^{(1)} = \mathbf{1}_{\mathcal{E}_{k+1}}$ .

Because of the last fact, viewing  $\mathcal{H}_{\mathcal{Y}_{k+1}}$  as being canonically isomorphic to  $\mathcal{H}_{\mathcal{Y}_k} \otimes \mathbb{C}^2$ , we have

$$H(\mathcal{G}_{k+1}, J_{k+1}^{(0)}) = H(\mathcal{G}_k, J_k) \otimes \mathbb{1}_{\mathbb{C}^2}.$$

We may choose a vector  $\phi \in \mathcal{H}_{\mathcal{Y}_k}(n, n)$  such that  $\|\phi\| = 1$  and  $H(\mathcal{G}_k, J_k)\phi = \mathfrak{E}_n(\mathcal{G}_k, J_k)\phi$ . Then we may take  $\psi = \phi \otimes |\uparrow\rangle$  in  $\mathcal{H}_{\mathcal{Y}_{k+1}}$ . This will be in  $\mathcal{H}_{\mathcal{Y}_{k+1}}(n, n)$ , and  $H(\mathcal{G}_{k+1}, J_{k+1}^{(0)})\psi = \mathfrak{E}_n(\mathcal{G}_k, J_k)\phi$ . These facts are non-trivial. But we leave them to the reader as an exercise in the definitions of magnetization and spin, and an application of “addition of angular momentum.” This proves that

$$\mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}^{(0)}) \leq \mathfrak{E}_n(\mathcal{G}_k, J_k). \quad (37)$$

There are two cases to consider.

In Case 1, we have the condition  $\mathfrak{E}_n(\mathcal{G}_{k+1}) \leq \mathfrak{E}_n(\mathcal{G}_k, J_k)$ . Then we take  $J_{k+1} = J_{k+1}^{(1)}$ . In other words, in that case, we choose  $t = 1$ . This means  $J_{k+1} = \mathbf{1}_{\mathcal{E}_{k+1}}$  in this case. That means that  $\mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}) = \mathfrak{E}_n(\mathcal{G}_{k+1})$ .

In Case 2 we have  $\mathfrak{E}_n(\mathcal{G}_{k+1}) > \mathfrak{E}_n(\mathcal{G}_k, J_k)$ . Then we know the following. The mapping  $t \mapsto \mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}^{(t)})$  is non-decreasing and continuous in  $t$ . At  $t = 0$  we get a number at most equal to  $\mathfrak{E}_n(\mathcal{G}_k, J_k)$  according to (37). At  $t = 1$  we get  $\mathfrak{E}_n(\mathcal{G}_{k+1})$ , a number strictly greater than  $\mathfrak{E}_n(\mathcal{G}_k, J_k)$ . Therefore, there is at least one intermediate value  $t \in [0, 1]$  with  $\mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}^{(t)}) = \mathfrak{E}_n(\mathcal{G}_k, J_k)$ . We let  $\mathcal{T}$  be the set of all such  $t$ 's, and we let  $t_* = \sup(\mathcal{T})$ . By continuity of the mapping,  $t_* \in \mathcal{T}$ . We let  $J_{k+1} = J_{k+1}^{(t_*)}$ . In this case, we have  $\mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}) = \mathfrak{E}_n(\mathcal{G}_k, J_k)$ .

In Case 1, we assumed  $\mathfrak{E}_n(\mathcal{G}_{k+1}) \leq \mathfrak{E}_n(\mathcal{G}_k, J_k)$  and obtained  $J_{k+1}$  such that  $\mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}) = \mathfrak{E}_n(\mathcal{G}_{k+1})$ . In Case 2, we assumed  $\mathfrak{E}_n(\mathcal{G}_{k+1}) > \mathfrak{E}_n(\mathcal{G}_k, J_k)$  and obtained  $J_{k+1}$  such that

$$\mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}) = \mathfrak{E}_n(\mathcal{G}_k, J_k).$$

Therefore, in either case, we have

$$\mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}) = \min\{\mathfrak{E}_n(\mathcal{G}_k, J_k), \mathfrak{E}_n(\mathcal{G}_{k+1})\}.$$

By the induction hypothesis and (36) we see that in fact

$$\mathfrak{E}_n(\mathcal{G}_{k+1}, J_{k+1}) = \min\{\mathfrak{E}_n(\mathcal{G}_{2n}), \dots, \mathfrak{E}_n(\mathcal{G}_{k+1})\}.$$

But this is precisely (36) with  $k$  replaced by  $k+1$ . So the induction step is proved.  $\square$

**Proof of Proposition 3.3:** By Lemma 5.8 we know that there exists a diluted system  $J_{2n}, \dots, J_N$  such that  $H(\mathcal{G}_N, J_N) = H(\mathcal{G}_N)$ . Then, by Corollary 5.6, we see that  $(\mathcal{G}_k, J_k)$  satisfies FOEL- $n$ , for each  $k = 2n, \dots, N$ . Therefore, we have, for any  $k \in \{2n, \dots, N\}$ ,

$$\min_{r \geq n} \mathfrak{E}_r(\mathcal{G}_k, J_k) \geq \mathfrak{E}_n(\mathcal{G}_k, J_k).$$

But since  $(\mathfrak{E}_n(\mathcal{G}_k, J_k))_{k=2n}^N$  is non-increasing, this means

$$\min_{r \geq n} \mathfrak{E}_r(\mathcal{G}_k, J_k) \geq \mathfrak{E}_n(\mathcal{G}_k, J_k) \geq \mathfrak{E}_n(\mathcal{G}_N, J_N) = \mathfrak{E}_n(\mathcal{G}_N), \quad (38)$$

where the last equality holds because  $H(\mathcal{G}_N, J_N) = H(\mathcal{G}_N)$ . Moreover, since  $J_k \leq \mathbf{1}_{\mathcal{E}_k}$ , we know that as operators  $H(\mathcal{G}_k, J_k) \leq H(\mathcal{G}_k)$ , meaning the difference is positive semi-definite. This is the type of ordering which leads to inequalities for the Rayleigh quotients. Therefore, we also have  $\mathfrak{E}_r(\mathcal{G}_k, J_k) \leq \mathfrak{E}_r(\mathcal{G}_k)$  for all  $r$ . Putting this together with (38), we obtain

$$\mathfrak{E}_n(\mathcal{G}_N) \leq \mathfrak{E}_r(\mathcal{G}_k, J_k) \leq \mathfrak{E}_r(\mathcal{G}_k),$$

for each  $k \in \{2n, \dots, N\}$  and each  $r \in \{n, \dots, \lfloor k/2 \rfloor\}$ .

That is what we wanted to prove.  $\square$

## 6 Proof of Key Step II – Part A: Linear Spin Wave Approximation

We now wish to prove Key Step II, meaning Theorem 3.10. We also needed Proposition 3.9 in the conditional proof of the main theorem. But we will prove Proposition 3.9 in Appendix ?? . It will be easy, but it will follow a direct calculation using the explicit formula for the approximate eigenvectors.

Theorem 3.10 and Proposition 3.9 give an approximation for the low-energy spectrum of the ferromagnetic Heisenberg Hamiltonian  $H_{\Lambda(d,N)}$  in the limit  $N \rightarrow \infty$ . This approximation is part of the linear spin wave approximation: it is the specialization of the linear spin wave approximation at lowest energies, which are energies bounded by a finite multiple of the spectral gap. We will complete this part in several steps.

- In this section, Part A, we introduce the transformation needed to compare to an ideal Bose gas. We also state the main results about this transformation and an approximate inverse, showing that these transformations do not distort low-energy trial wave functions, too much. This relies on two important types of bounds.
- In Part B, we give variational arguments, combined with Chebyshev's inequality, to apply this comparison to prove Theorem 3.10. To apply Chebyshev's inequality, we will also need Proposition 3.9. But this proposition will be proved in Appendix ??
- In Part C, we prove one of the two important types of bounds, which is a trace-theorem type bound. It shows that, relative to the ideal Bose gas, a variational trial wave vector with low energy does not lose much norm in the contraction transformation to the quantum spin system.
- In Part D, which is an extension theorem, we prove a complementary result: given a low energy trial wave vector for the quantum spin system, we may extend it to a wave vector for the ideal Bose gas such that we raise the energy only by a bounded factor.

The linear spin wave approximation goes back to Bloch [3], who used the unproven hypothesis to derive a non-rigorous formula for the low temperature pressure. There were further developments by Dyson in [14, 15] who showed there must be corrections to the naive approach. The first rigorous analysis was initiated by Conlon and Solovej who obtained one-sided bounds for the pressure [7, 8]. The bounds were improved by Balint Toth [30]. Importantly, Toth also introduced a new representation for the quantum Heisenberg ferromagnet in terms of the interchange process<sup>1</sup>.

<sup>1</sup>The interchange process is also known as the stirring process in [21], where it is discussed especially as the graphical representation for the symmetric exclusion process. Also, coincidentally, Larry Thomas had noted a relation between the Heisenberg model and the symmetric exclusion process in [29], although he did not capitalize on this observation.

More recently, the linear spin wave approximation was reconsidered by Correggi, Giuliani and Seiringer who improved Conlon and Solovej's and Toth's bounds, to verify Bloch's original ansatz for the asymptotic formula for the free energy at low temperatures [11]. Their article is a good reference for the linear spin wave approximation.

For  $n \in \{0, 1, \dots\}$ , let  $\ell^2(\mathcal{V}^n)$  denote the vector space (of dimension  $|\mathcal{V}|^n$ ) of functions  $F : \mathcal{V}^n \rightarrow \mathbb{C}$ , with the usual  $\ell^2$  norm  $\|F\|^2 = \sum_{(x_1, \dots, x_n) \in \mathcal{V}^n} |F(x_1, \dots, x_n)|^2$ . Then, for each  $k \in \{1, \dots, n\}$ , we denote the  $k$ th particle graph Laplacian as  $\mathfrak{H}_{\mathcal{G},k}^{(n)} : \ell^2(\mathcal{V}^n) \rightarrow \ell^2(\mathcal{V}^n)$ , where, for each  $F \in \ell^2(\mathcal{V}^n)$  and each  $(x_1, \dots, x_n) \in \mathcal{V}^n$ ,

$$\mathfrak{H}_{\mathcal{G},k}^{(n)} F(x_1, \dots, x_n) = \frac{1}{2} \sum_{y \in \mathcal{N}(\mathcal{E}, x_k)} (F(x_1, \dots, x_n) - F(x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_n)). \quad (39)$$

Given any  $\pi \in S_n$ , we may define the operator  $U_{\mathcal{V}}(\pi) : \ell^2(\mathcal{V}^n) \rightarrow \ell^2(\mathcal{V}^n)$  such that

$$U_{\mathcal{V}}(\pi) F(x_1, \dots, x_n) = F(x_{\pi^{-1}(1)}, \dots, x_{\pi^{-1}(n)}),$$

for each  $F \in \ell^2(\mathcal{V}^n)$  and each  $(x_1, \dots, x_n) \in \mathcal{V}^n$ .

**Lemma 6.1** (a) For each  $\pi \in S_n$ , the operator  $U_{\mathcal{V}}(\pi)$  is unitary.

(b) For each  $k \in \{1, \dots, n\}$  and each  $\pi \in S_n$ ,

$$U_{\mathcal{V}}(\pi)^* \mathfrak{H}_{\mathcal{G},k}^{(n)} U_{\mathcal{V}}(\pi) = \mathfrak{H}_{\mathcal{G},\pi^{-1}(k)}^{(n)}.$$

We will prove this easy lemma in Appendix C.

The total  $n$ -particle Hamiltonian for  $n$  free particles on  $\mathcal{G}$  is  $\mathfrak{H}_{\mathcal{G}}^{(n)} : \ell^2(\mathcal{V}^n) \rightarrow \ell^2(\mathcal{V}^n)$ , given by

$$\mathfrak{H}_{\mathcal{G}}^{(n)} = \sum_{k=1}^n \mathfrak{H}_{\mathcal{G},k}^{(n)}.$$

Since this is the uniform sum,  $\mathfrak{H}_{\mathcal{G}}^{(n)}$  commutes with  $U_{\mathcal{V}}(\pi)$  for each  $\pi \in S_n$ . We denote the symmetrization projection as  $\mathfrak{S}_{\mathcal{V}}^{(n)} : \ell^2(\mathcal{V}^n) \rightarrow \ell^2(\mathcal{V}^n)$ ,

$$\mathfrak{S}_{\mathcal{V}}^{(n)} = \frac{1}{n!} \sum_{\pi \in S_n} U_{\mathcal{V},\pi}.$$

Hence,  $\mathfrak{H}_{\mathcal{G}}^{(n)}$  commutes with  $\mathfrak{S}_{\mathcal{V}}^{(n)}$ .

**Remark 6.2** Given the graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , we may define two different graph structures on  $\mathcal{V}^n$ . Let  $\Phi_n(\mathcal{E})$  denote the set of all two-element sets  $\{(x_1, \dots, x_n), (x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_n)\}$  for all possible choices of  $(x_1, \dots, x_n) \in \mathcal{V}^n$ ,  $k \in \{1, \dots, n\}$  and  $y \in \mathcal{N}(\mathcal{E}, x_k)$ . Then the graph Laplacian for  $(\mathcal{V}^n, \Phi_n(\mathcal{E}))$  is  $\mathfrak{H}_{\mathcal{G}}^{(n)}$ , using the definition as in (8).

Considering the remark, there is another important edge set. Let

$$\mathcal{I}_n(\mathcal{V}) = \{(x_1, \dots, x_n) \in \mathcal{V}^n : \exists i, j \in \{1, \dots, n\} \text{ s.t. } i \neq j \text{ and } x_i = x_j\}. \quad (40)$$

Then we may define  $\Theta_n(\mathcal{E})$  to denote the set of all edges in  $\Phi_n(\mathcal{E})$  satisfying the additional condition that neither endpoint is in  $\mathcal{I}_n(\mathcal{V})$ . Then the graph  $(\mathcal{V}^n \setminus \mathcal{I}_n(\mathcal{V}), \Theta_n(\mathcal{V}))$  has a Laplacian closely connected  $H_{\mathcal{G}} \upharpoonright \mathcal{H}_{\mathcal{V}}^{\text{mag}}(n)$ .

If we restrict  $\mathfrak{H}_{\mathcal{G}}^{(n)}$  to  $\mathfrak{S}_{\mathcal{V}}^{(n)}(\ell^2(\mathcal{V}^n))$ , then this is the Hamiltonian for  $n$  particles in an ideal Bose gas on  $\mathcal{G}$ . Part of the linear spin wave approximation states that this is a “good” approximation for the ferromagnetic Heisenberg Hamiltonian  $H_{\mathcal{G}}$  restricted to  $\mathcal{H}_{\mathcal{V}}^{\text{mag}}(n)$ , in a certain sense.

Recall the mapping  $T_{\mathcal{V}}^{(n)} : \ell^2(\mathcal{V}^n) \rightarrow \mathcal{H}_{\mathcal{V}}$ , from Definition 3.5. The following is easy to prove.

**Lemma 6.3** (a) For each  $n \in \{1, \dots, |\mathcal{V}|\}$ , and each  $F \in \ell^2(\mathcal{V}^n)$ ,

$$\|T_{\mathcal{V}}^{(n)} F\| \leq \|F\|.$$

(b) For each  $n \in \{1, \dots, |\mathcal{V}|\}$ , the operator  $\mathfrak{H}_{\mathcal{G}}^{(n)}$  on  $\ell^2(\mathcal{V}^n)$  is positive semi-definite, and

$$T_{\mathcal{V}}^{(n)} \mathfrak{H}_{\mathcal{G}}^{(n)} (T_{\mathcal{V}}^{(n)})^* \geq (H_{\mathcal{G}} \upharpoonright \mathcal{H}_{\mathcal{V}}^{\text{mag}}(n)).$$

We will also prove this easy result in Appendix C, since it follows a similar type of argument as the proof of Lemma 6.1. Part (b) will follow from the fact already expressed in Remark 6.2, as well as the fact, which will be used several more times, that the mapping from  $\mathcal{G}$  to  $-\Delta_{\mathcal{G}}$ , as defined in equation (8), is increasing, relative to the cone of positive semi-definite operators, as we increase edges. (This was also a key to the arguments in Section 5.) When we speak about this fact, now, we are usually thinking of the edge set as being  $\Phi_n(\mathcal{E})$  or  $\Theta_n(\mathcal{E})$ , as opposed to the edge set of the original graph.

To prove a part of the linear spin wave approximation, we need some complementary results, as well. For these inequalities, we revert to considering the special family of graphs from before.

**Definition 6.4** Fix,  $d, n \in \{1, 2, \dots\}$ . For each  $N \in \{1, 2, \dots\}$ , define  $\tilde{T}_{d,N}^{(n)} : \ell^2(\mathbb{B}^d(L^+(d, N))) \rightarrow \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n)$  such that for any  $F \in \ell^2(\mathbb{B}^d(L^+(d, N)))$ , we first take the restriction  $F \upharpoonright (\Lambda(d, N))^n$ , restricting to the domain of  $T_{\Lambda(d,N)}^{(n)}$ , and then we take  $\tilde{T}_{d,N}^{(n)} F = T_{\Lambda(d,N)}^{(n)} (F \upharpoonright (\Lambda(d, N))^n)$ .

A trivial corollary of Lemma 6.3 is this

**Lemma 6.5** For any  $d, L \in \{1, 2, \dots\}$  and  $n \in \{0, 1, \dots\}$ , we have that  $\|\tilde{T}_{d,N}^{(n)}\| \leq 1$  and

$$\tilde{T}_{d,N}^{(n)} \mathfrak{H}_{\mathbb{B}^d(L^+(d,N))}^{(n)} (\tilde{T}_{d,N}^{(n)})^* \geq (H_{\Lambda(d,N)} \upharpoonright \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n)). \quad (41)$$

Once again, we will prove this simple lemma in Appendix C. But the idea is just to use the fact, again, that increasing edges of a graph  $\mathcal{G}$  increases the graph Laplacian  $-\Delta_{\mathcal{G}}$ , relative to the psd cone. There are clearly at least as many edges in  $\mathbb{B}^d(L^+(d, N))$  as in  $\Lambda(d, N)$  since  $\Lambda(d, N) \subseteq \mathbb{B}^d(L^+(d, N))$ .

**Proposition 6.6** Fix,  $d, n \in \{1, 2, \dots\}$ . Then there is an  $N_1(d, n) \in \{1, 2, \dots\}$  and two constants  $C_1(d, n)$ ,  $C_2(d, n)$  such that, for any  $N \geq N_1(d, n)$  and for any  $F \in \mathfrak{S}_{\mathbb{B}^d(L^+(d,N))}^{(n)}(\ell^2(\mathbb{B}^d(L^+(d, N))^n))$ , we have

$$\|\tilde{T}_{d,N}^{(n)} F\|^2 \geq \left(1 - \frac{C_1(d, n)}{L^+(d, N)}\right) \|F\|^2 - C_2(d, N) L^+(d, N) \langle F, \mathfrak{H}_{\Lambda(d,N)}^{(n)} F \rangle. \quad (42)$$

We will prove Proposition 6.6 in Section 8.

**Proposition 6.7** Fix  $d, n \in \{1, 2, \dots\}$ . There exists an  $N_2(d, n) \in \{1, 2, \dots\}$  and a constant  $C_3(d, n)$  such that, for each  $N \in \{1, 2, \dots\}$  satisfying  $N \geq N_2(d, n)$ , there is a linear mapping

$$\Xi_{d,N}^{(n)} : \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n) \rightarrow \mathfrak{S}_{\mathbb{B}^d(L^+(d,N))}^{(n)} \left( \ell^2 \left( (\mathbb{B}^d(L^+(d, N)))^n \right) \right),$$

such that  $\tilde{T}_{d,N}^{(n)} \cdot \Xi_{d,N}^{(n)}$  is the identity on  $\mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n)$  and such that

$$\langle \Xi_{d,N}^{(n)} \Psi, \mathfrak{H}_{\mathbb{B}^d(L^+(d,N))}^{(n)} \Xi_{d,N}^{(n)} \Psi \rangle \leq C_3(d, n) \langle \Psi, H_{\Lambda(d,N)} \Psi \rangle, \quad (43)$$

for each  $\Psi \in \mathcal{H}_{\Lambda(d,N)}^{\text{mag}}(n)$ .

We will prove this proposition in Section 9.

In the next section, we will explain how to prove Theorem 3.10. We do thi by combining Proposition 3.9, Proposition 6.6, Proposition prop:extension0, the Rayleigh-Ritz variational approach and Chebyshev's inequality.

## 7 Proof of Key Step II – Part B: Variational Argument

Let us begin by noting the formula for the spectrum of the non-interacting Bose gas.

**Definition 7.1** For  $d, L \in \{1, 2, \dots\}$ , let  $\mathcal{O}(d, L)$  denote the set of all “occupation functions,” defined to be functions  $\nu : \{0, \dots, L-1\}^d \rightarrow \{0, 1, \dots\}$ . For  $n \in \{0, 1, \dots\}$  let  $\mathcal{O}_n(d, L)$  be the set of  $\nu \in \mathcal{O}(d, L)$  such that

$$\sum_{\kappa \in \{0, \dots, L-1\}^d} \nu(\kappa) = n.$$

For each choice of  $\nu \in \mathcal{O}_n(d, L)$ , let us define  $\mathfrak{K}_{d,L}(\nu)$  to be a subset of  $(\{0, \dots, L-1\}^d)^n$ , defined as the set of all  $(\kappa_1, \dots, \kappa_n)$  such that

$$\forall \kappa \in \{0, \dots, L-1\}^d : |\{k \in \{1, \dots, n\} : \kappa_k = \kappa\}| = \nu(\kappa).$$

**Definition 7.2** For each  $\nu \in \mathcal{O}_n(d, L)$ , define a function  $\tilde{F}_{d,L}^{(n)}(\nu; \cdot) \in \ell^2((\mathbb{B}^d(L))^n)$  by the formula

$$\tilde{F}_{d,L}^{(n)}(\nu; \mathbf{r}_1, \dots, \mathbf{r}_n) = L^{-nd/2} |\mathfrak{K}_{d,L}(\nu)|^{-1/2} \sum_{(\kappa_1, \dots, \kappa_n) \in \mathfrak{K}_{d,L}(\nu)} \prod_{k=1}^n \prod_{j=1}^d f(L^{-1} \kappa_{k,j}, r_{k,j}), \quad (44)$$

for each  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in (\mathbb{B}^d(L))^n$ , where  $f(\xi, r)$  for  $\xi \in \mathbb{R}$  and  $r \in \mathbb{Z}$  is given by Definition 3.6.

The spectrum of  $\mathfrak{H}_{\mathbb{B}^d(L)}^{(n)}$  restricted to the invariant subspace  $\mathfrak{S}_{\mathbb{B}^d(L)}^{(n)}(\ell^2((\mathbb{B}^d(L))^n))$  may then be summarized, as follows.

**Lemma 7.3** For  $d, L \in \{1, 2, \dots\}$  and  $n \in \{0, 1, \dots\}$ , an orthonormal basis of  $\mathfrak{S}_{\mathbb{B}^d(L)}^{(n)}(\ell^2((\mathbb{B}^d(L))^n))$  is  $(\tilde{F}_{d,L}^{(n)}(\nu; \cdot) : \nu \in \mathcal{O}_n(d, L))$ , and

$$\mathfrak{H}_{\mathbb{B}^d(L)}^{(n)} \tilde{F}_{d,L}^{(n)}(\nu; \cdot) = \sum_{\kappa \in \{0, \dots, L-1\}^d} \nu(\kappa) \sum_{j=1}^d 2 \sin^2 \left( \frac{\pi \kappa_j}{2L} \right) \tilde{F}_{d,L}^{(n)}(\nu; \cdot), \quad (45)$$

for each  $\nu \in \mathcal{O}_n(d, L)$ .

This easy lemma will also be proved in Appendix B. This calculation is well-known, since it is the spectrum of a non-interacting ideal Bose gas on a finite box in a lattice. But we will include it for completeness.

We may now prove part of Proposition 3.9. First, given  $(\kappa_1, \dots, \kappa_n) \in \mathcal{K}(d, n)$ , if  $L$  is sufficiently large that  $\kappa_k \in \{0, \dots, L-1\}^d$  for each  $k \in \{1, \dots, n\}$ , then let us define  $\nu_{(\kappa_1, \dots, \kappa_n)}$  to be the associated point of  $\mathcal{O}_n(d, L)$ :

$$\nu_{(\kappa_1, \dots, \kappa_n)}(\kappa) = \sum_{k=1}^n \mathbf{1}_{\{\kappa_k = \kappa\}}, \quad (46)$$

for each  $\kappa \in \{0, \dots, L-1\}^d$ . A calculation then shows

$$\tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n) = n! |\mathfrak{K}_{d,L^+(d,N)}(\nu_{(\kappa_1, \dots, \kappa_n)})|^{-1/2} \tilde{T}_{d,N}^{(n)} \tilde{F}_{d,L^+(d,N)}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \cdot). \quad (47)$$

Next we want to establish variational upper bounds, and then lower bounds using Chebyshev's inequality. For the variational upper bounds we will use the following: For a self-adjoint operator  $A$  on a  $d$ -dimensional Hilbert space  $\mathcal{H}$ , we have that

$$\lambda_k = \min \left\{ \max_{v \in \text{span}(\{v_1, \dots, v_k\}) \setminus \{0\}} \frac{\langle v, Av \rangle}{\|v\|^2} : v_1, \dots, v_k \in \mathcal{H} \text{ are linearly independent} \right\}, \quad (48)$$

for each  $k \in \{1, \dots, d\}$ , where the eigenvalues of  $A$  are enumerated in non-decreasing order  $\lambda_1 \leq \dots \leq \lambda_d$ , repeating according to multiplicity. With this, we may prove the first part of the arguments.

**Lemma 7.4** *Let  $\mathcal{S}$  be a finite subset of  $\tilde{\mathcal{K}}(d, n)$ . Let  $|\mathcal{S}|$  denote the cardinality of  $\mathcal{S}$ , as usual. But also let us introduce a new notation*

$$\|\mathcal{S}\|^2 = \max \left\{ \sum_{k=1}^n \sum_{j=1}^d \kappa_{k,j}^2 : [(\kappa_1, \dots, \kappa_n)]_{S_n} \in \mathcal{S} \right\}.$$

(The sum is permutation independent, so the definition does not depend on which representatives of each equivalence class are chosen.) Recall the definition of the spectral subspaces from (14). For each  $\epsilon > 0$ , there exists an  $N_0(\mathcal{S}, \epsilon)$  such that

$$\dim(\mathfrak{L}_{\Lambda(d, N)}(0, \gamma \cdot [L^+(d, N)]^{-2}(\|\mathcal{S}\|^2 + \epsilon))) \geq |\mathcal{S}|.$$

**Proof:** We consider the associated vectors  $\tilde{\Psi}_{d, N}^{(n)}(\kappa_1, \dots, \kappa_n)$ , one chosen for each  $[(\kappa_1, \dots, \kappa_n)]_{S_n} \in \mathcal{S}$  (and noting that the definition is permutation independent, so that the choice of which representative we choose is not important). By equation (21), we know they are linearly independent, for sufficiently large  $N$ . For any element of the span, we have some  $c : \mathcal{S} \rightarrow \mathbb{C}$  such that the vector is

$$\Phi = \sum_{[(\kappa_1, \dots, \kappa_n)]_{S_n} \in \mathcal{S}} c([(\kappa_1, \dots, \kappa_n)]_{S_n}) \tilde{\Psi}_{d, N}^{(n)}(\kappa_1, \dots, \kappa_n)$$

Let us also define

$$G = \sum_{[(\kappa_1, \dots, \kappa_n)]_{S_n} \in \mathcal{S}} c([(\kappa_1, \dots, \kappa_n)]_{S_n}) \tilde{F}_{d, L^+(d, N)}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \cdot).$$

Then, using equation (21) from Proposition 3.9, that for any  $\epsilon > 0$  and sufficiently large  $N$  we have

$$1 \geq \frac{\|\Psi\|^2}{\sum_{[(\kappa_1, \dots, \kappa_n)]_{S_n} \in \mathcal{S}} |c([(\kappa_1, \dots, \kappa_n)]_{S_n})|^2 (n!)^2 |\mathfrak{K}_{d, L^+(d, N)}(\nu)|^{-1}} \geq 1 - \epsilon, \quad (49)$$

The denominator in the middle expression is equal to  $\|G\|^2$ . The bounds in (49) are both upper and lower bounds for the denominator term in (48).

But for the numerator, we have just upper bounds, so far. Using (41) in Lemma 6.5, and using (47), we have

$$\begin{aligned} \langle \Psi, H_{\Lambda(d, N)} \Psi \rangle &\leq \langle G, \mathfrak{H}_{\mathbb{B}^d(L^+(d, N))}^{(n)} G \rangle \\ &= \sum_{[(\kappa_1, \dots, \kappa_n)]_{S_n} \in \mathcal{S}} |c([(\kappa_1, \dots, \kappa_n)]_{S_n})|^2 (n!)^2 |\mathfrak{K}_{d, L^+(d, N)}(\nu)|^{-1} \sum_{k=1}^n \sum_{j=1}^d \kappa_{k,j}^2. \end{aligned}$$

Then the result follows by a max-norm bound of the final sum on the right-hand-side using  $\|\mathcal{S}\|^2$  to replace  $\sum_{k=1}^n \sum_{j=1}^d \kappa_{k,j}^2$  (and (49) for the denominator).  $\square$

Now we may complete the proof of Theorem 3.10.

**Proof of Theorem 3.10:** We now use Proposition 6.7. Suppose that  $m$  is fixed, as well as a small parameter  $\epsilon \in (0, 1/2)$ , and that we have  $R$  orthonormal eigenvectors of  $H_{\Lambda(d,N)}$ , called  $\Phi_1, \dots, \Phi_R \in \mathfrak{S}_{\mathbb{B}^d(L^+(d,N))}^{(n)}(\ell^2((\mathbb{B}^d(L^+(d,N)))^n))$ , with energy eigenvalues equal to  $\gamma \cdot [L^+(d,N)]^{-2} \lambda_r$  for each  $r \in \{1, \dots, R\}$ . And suppose that the rescaled eigenvalues satisfy  $\lambda_1, \dots, \lambda_R \in [0, m + \epsilon]$ . Let  $G_r = \Xi_{d,N}^{(n)} \Phi_r$  for each  $r \in \{1, \dots, R\}$ . Then by Proposition 6.7, we know that

$$\langle G_r, \mathfrak{H}_{\mathbb{B}^d(L^+(d,N))}^{(n)} G_r \rangle \leq \gamma \cdot [L^+(d,N)]^{-2} (m + \epsilon) C_3(d, n), \quad (50)$$

where  $C_3(d, n)$  is a large but fixed constant, independent of  $N$ . We also know that  $\tilde{T}_{d,N}^{(n)} G_r = \Phi_r$  for each  $r$ . In particular, this means that

$$\|G_r\|^2 \geq \|\tilde{T}_{d,N}^{(n)} G_r\|^2 = \|\Phi_r\|^2 = 1,$$

by equation (41). Then, using Lemma 7.3 and Chebyshev's inequality, we may write

$$G_r = \sum_{\nu \in \mathcal{O}_n(d, L^+(d, N))} c_r(\nu) \tilde{F}_{d, L^+(d, N)}^{(n)}(\nu; \cdot),$$

and we may bound the coefficients of  $\nu$ 's for which  $\tilde{F}_{d, L^+(d, N)}^{(n)}(\nu; \cdot)$  has a large eigenvalue. More precisely, let us use the bound  $\sin(\theta) \geq 2\theta/\pi$  for  $\theta \in [0, \pi/2]$ , and let us define, for a given  $M$ ,

$$\mathcal{S}_M = \left\{ \nu \in \mathcal{O}_n(d, L^+(d, N)) : \sum_{\kappa \in \{0, \dots, L-1\}^d} \nu(\kappa) \sum_{j=1}^d \kappa_j^2 \leq M \right\}.$$

Then we can say that

$$\sum_{\nu \in \mathcal{O}_n(d, L^+(d, N)) \setminus \mathcal{S}_M} |c_r(\nu)|^2 \leq \frac{\gamma \cdot [L^+(d, N)]^{-2} (m + \epsilon) C_3(d, n)}{2[L^+(d, N)]^{-2} M} = \frac{\gamma(m + \epsilon) C_3(d, n)}{2M}.$$

Let us define the truncated function, where we cut-off the coefficients in energy modes higher than the natural cutoff related to  $M$ :

$$\tilde{G}_{r,M} = \sum_{\nu \in \mathcal{S}_M} c_r(\nu) \tilde{F}_{d, L^+(d, N)}^{(n)}(\nu; \cdot).$$

Then we have shown that  $\|G_r - \tilde{G}_{r,M}\|^2 \leq C'_3(d, n)(m + \epsilon)/M$  for some constant  $C'_3(d, n)$ . Moreover, by (41), again, this implies

$$\|\Phi_r - \tilde{T}_{d,N}^{(n)} \tilde{G}_{r,M}\|^2 = \|\tilde{T}_{d,N}^{(n)} (G_r - \tilde{G}_{r,M})\|^2 \leq \|G_r - \tilde{G}_{r,M}\|^2 \leq \frac{C'_3(d, n)(m + \epsilon)}{M}.$$

So we first choose  $M$  to be a sufficiently large, but fixed, multiple of  $m$  so that this fraction is small. In particular, then we may assume that the vectors  $\tilde{T}_{d,N}^{(n)} \tilde{G}_{r,M}$  are approximately orthonormal, where each inner product deviates from  $\delta_{r,s}$  by an amount not larger than  $O((m + \epsilon)/M)$ . Then we notice that  $\mathcal{S}_M$  is still finite, and in particular the energy of  $\tilde{F}_{d, L^+(d, N)}^{(n)}(\nu; \cdot)$  for any  $\nu \in \mathcal{S}_M$  is actually

bounded by  $\gamma \cdot [L^+(d, N)]^{-2}M$ . So, if we assume that  $N$  is large, then  $\tilde{T}_{d,N}^{(n)}$  does not deform each  $\tilde{G}_r$  by very much. Another way of saying this is as follows. We have

$$\tilde{T}_{d,N}^{(n)}\tilde{G}_r = \sum_{\nu \in \mathcal{S}_M} c_r(\nu) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)),$$

where we pick one element  $(\kappa_1(\nu), \dots, \kappa_n(\nu)) \in \mathfrak{K}_{d,L^+(d,N)}(\nu)$  for each  $\nu \in \mathcal{S}_M$ .

Now, using the fact that  $\Phi_r$  is an exact eigenvector, we have, for each  $\nu \in \mathcal{S}_M$ ,

$$\begin{aligned} \lambda_r \langle \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)), \Phi_r \rangle &= \langle \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)), \gamma^{-1} L^2 H_{\Lambda(d,N)} \Phi_r \rangle \\ &= \langle \gamma^{-1} L^2 H_{\Lambda(d,N)} \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)), \Phi_r \rangle \\ &= \|\nu\|^2 \cdot \langle \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)), \Phi_r \rangle \\ &\quad + \langle (\gamma^{-1} L^2 H_{\Lambda(d,N)} - \|\nu\|^2) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)), \Phi_r \rangle. \end{aligned}$$

So we see that, assuming  $\|\nu\|^2 \neq \lambda_r$ , we have

$$\begin{aligned} |\langle \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)), \Phi_r \rangle|^2 &\leq \frac{\left| \langle (\gamma^{-1} L^2 H_{\Lambda(d,N)} - \|\nu\|^2) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)), \Phi_r \rangle \right|^2}{|\lambda_r - \|\nu\|^2|^2} \\ &\leq \frac{\|(\gamma^{-1} L^2 H_{\Lambda(d,N)} - \|\nu\|^2) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu))\|^2}{|\lambda_r - \|\nu\|^2|^2}, \end{aligned}$$

by Cauchy-Schwarz. But we have already established in Proposition 3.9 that the numerator is vanishingly small as  $L \rightarrow \infty$ . Therefore, since  $\lambda_r < m + \epsilon$ , if  $\nu \in \mathcal{S}_M \setminus \mathcal{S}_m$ , then we can establish that  $|\langle \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)), \Phi_r \rangle|^2$  is small.

But since  $\Phi_r$  is close to  $\tilde{T}_{d,N}^{(n)}\tilde{G}_r$  in norm, this implies that  $c_r(\nu)$  is small for such  $\nu$ 's. So, defining

$$\tilde{\Phi}_{r,m} = \sum_{\nu \in \mathcal{S}_m} c_r(\nu) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1(\nu), \dots, \kappa_n(\nu)),$$

we have that  $\|\Phi_r - \tilde{\Phi}_r\|$  is small for each  $r$ . But  $\Phi_1, \dots, \Phi_R$  are orthonormal, while each  $\tilde{\Phi}_r$  is in a subspace spanned by  $|\mathcal{S}_m|$  vectors. Therefore, this requires  $R \leq |\mathcal{S}_m|$ . Combined with Lemma 7.4, this establishes the correct dimensions for the spectral subspaces. By Proposition 3.9, we have the correct number of trial wavefunctions which are also approximate eigenvectors. Once this is established, the remainder of the proof immediately follows.  $\square$

## 8 Proof of Key Step II – Part C: Trace theorem type bound

In this section we prove Proposition 6.6. We consider  $F \in \mathfrak{S}_{\mathbb{B}^d(L^+(d,N))}^{(n)}(\ell^2(\mathbb{B}^d(L^+(d,N))^n))$ . Let us allow ourselves to abbreviate  $L^+(d, N)$  as  $L$ . Then

$$\|F\|^2 - \|\tilde{T}_{d,N}^{(n)}F\|^2 = \sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n} (1 - \mathbf{1}_{\Lambda(d,N)^n \setminus \mathcal{J}_n(\Lambda(d,N))}(\mathbf{r}_1, \dots, \mathbf{r}_n)) |F(\mathbf{r}_1, \dots, \mathbf{r}_n)|^2.$$

Note that the first factor in the summation selects only those points  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n$  in the complement of  $\Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ . This set is the union of  $\mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n$  and  $\mathcal{J}_n(\Lambda(d, N))$ . So

$$\|F\|^2 - \|\tilde{T}_{d,N}^{(n)}F\|^2 = \sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n} |F(\mathbf{r}_1, \dots, \mathbf{r}_n)|^2 + \sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathcal{J}_n(\Lambda(d, N))} |F(\mathbf{r}_1, \dots, \mathbf{r}_n)|^2.$$



We will consider these two sums, separately. But both will use the same elementary inequality which is a discrete version of the trace theorem inequality (as it is called in the theory of PDE's):

**Lemma 8.1** *Suppose that  $f : \{1, \dots, L\} \rightarrow \mathbb{C}$  is a function. Then*

$$|f(1)|^2 \leq \frac{2L}{3} \sum_{\ell=1}^{L-1} |f(\ell) - f(\ell+1)|^2 + \frac{2(L-1)}{L^2} \sum_{\ell=1}^L |f(\ell)|^2. \quad (51)$$

**Proof:** Let  $\phi(\ell) = (L - \ell)/(L - 1)$ . Then, we have

$$f(1) = f(1)\phi(1) - f(L)\phi(L) = \sum_{\ell=1}^{L-1} [f(\ell)\phi(\ell) - f(\ell+1)\phi(\ell+1)].$$

Then we may rewrite

$$\begin{aligned} f(\ell)\phi(\ell) - f(\ell+1)\phi(\ell+1) &= \phi(\ell)[f(\ell) - f(\ell+1)] + [\phi(\ell) - \phi(\ell+1)]f(\ell+1) \\ &= \phi(\ell)[f(\ell) - f(\ell+1)] + \frac{1}{L-1} f(\ell+1). \end{aligned}$$

So

$$f(1) = \sum_{\ell=1}^{L-1} \phi(\ell)[f(\ell) - f(\ell+1)] + \frac{1}{L-1} \sum_{\ell=1}^{L-1} f(\ell+1).$$

Note that we can make the second sum slightly more symmetric, by including  $f(1)$  to obtain

$$\frac{L}{L-1} f(1) = \sum_{\ell=1}^{L-1} \phi(\ell)[f(\ell) - f(\ell+1)] + \frac{1}{L-1} \sum_{\ell=1}^L f(\ell).$$

Then, by the triangle inequality and Cauchy-Schwarz,

$$\frac{L}{L-1} |f(1)| \leq \left[ \sum_{\ell=1}^{L-1} \phi(\ell)^2 \right]^{1/2} \left[ \sum_{\ell=1}^{L-1} |f(\ell) - f(\ell+1)|^2 \right]^{1/2} + \frac{1}{\sqrt{L-1}} \left[ \sum_{\ell=1}^L |f(\ell)|^2 \right]^{1/2}.$$

Note that  $\sum_{\ell=1}^{L-1} \phi(\ell)^2 = (L-1)^{-2} \sum_{\ell=1}^{L-1} \ell^2$  and performing the sum, this is bounded by  $L^3/[3(L-1)^2]$ . Squaring and using the inequality  $(a+b)^2 \leq 2a^2 + 2b^2$  gives the result.  $\square$

Now we return to the problem of bounding

$$\sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n} |F(\mathbf{r}_1, \dots, \mathbf{r}_n)|^2, \quad \text{and} \quad \sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathcal{I}_n(\Lambda(d, N))} |F(\mathbf{r}_1, \dots, \mathbf{r}_n)|^2.$$

Given a point  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n$ , there is at least one  $k_0$  such that  $\mathbf{r}_{k_0} \in \mathbb{B}^d(L) \setminus \Lambda(d, N)$ . (We choose  $k_0$  to be the minimal element in case the number of possible choices is greater than 1.) Then, there is also some  $j_0 \in \{1, \dots, d\}$  such that  $r_{k_0, j_0} = L$ . (Again, in case of multiple choices we choose the minimal one.) Now we will consider a chain of points  $(\tilde{\mathbf{r}}_1(\ell), \dots, \tilde{\mathbf{r}}_n(\ell))$  for  $\ell \in \{0, \dots, \lfloor L/2 \rfloor - 1\}$ , as follows. We let  $\tilde{r}_{k_0, j_0}(\ell) = L - \ell$  and  $\tilde{r}_{k, j}(\ell) = r_{k, j}$  for all  $(k, j) \neq (k_0, j_0)$ .

Let us actually refer to these points as  $(\tilde{\mathbf{r}}_1(\ell; \mathbf{r}_1, \dots, \mathbf{r}_n), \dots, \tilde{\mathbf{r}}_n(\ell; \mathbf{r}_1, \dots, \mathbf{r}_n))$ , for later reference. By Lemma 8.1, we then know

$$\begin{aligned} & \sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n} |F(\mathbf{r}_1, \dots, \mathbf{r}_n)|^2 \\ & \leq \frac{L}{3} \sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n} \sum_{\ell=1}^{\lfloor L/2 \rfloor - 1} |F(\tilde{\mathbf{r}}_1(\ell; \mathbf{r}_1, \dots, \mathbf{r}_n)) - F(\tilde{\mathbf{r}}_1(\ell-1; \mathbf{r}_1, \dots, \mathbf{r}_n))|^2 \\ & \quad + \frac{2}{\lfloor L/2 \rfloor} \sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n} \sum_{\ell=0}^{\lfloor L/2 \rfloor - 1} |F(\tilde{\mathbf{r}}_1(\ell; \mathbf{r}_1, \dots, \mathbf{r}_n))|^2 \end{aligned} \quad (52)$$

Now, let us count the number of times any given edge in  $\Phi_n(\mathbb{B}^d(L))$ , say going from  $(\mathbf{r}'_1, \dots, \mathbf{r}'_n)$  to  $(\mathbf{r}''_1, \dots, \mathbf{r}''_n)$ , will occur as an edge, from  $(\tilde{\mathbf{r}}_1(\ell; \mathbf{r}_1, \dots, \mathbf{r}_n))$  to  $(\tilde{\mathbf{r}}_1(\ell-1; \mathbf{r}_1, \dots, \mathbf{r}_n))$ , for some point  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n$ . Note that  $r'_{kj} = r''_{kj}$  except for one choice  $(k_0, j_0)$ . But then we may determine that we must have  $(\mathbf{r}_1, \dots, \mathbf{r}_n)$  being the point such that  $r_{k_0, j_0} = L$  and  $r_{kj} = r'_{kj}$  for all  $(k, j) \neq (k_0, j_0)$ . So it can only occur at most one time. Thus, since each  $(\mathbf{r}'_1, \dots, \mathbf{r}'_n)$  is an endpoint for  $2dn$  edges (at most), we deduce that

$$\sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n} |F(\mathbf{r}_1, \dots, \mathbf{r}_n)|^2 \leq \frac{2L}{3} \langle F, \mathfrak{H}_{\mathbb{B}^d(L)}^{(n)} F \rangle + \frac{4nd}{\lfloor L/2 \rfloor} \|F\|^2. \quad (53)$$

Let us choose this point to declare that we are proving the proposition

**Proof of Proposition 6.6:** All that remains is to bound the second sum

$$\sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathcal{J}_n(\Lambda(d, N))} |F(\mathbf{r}_1, \dots, \mathbf{r}_n)|^2.$$

We use a similar argument as before. If  $(\mathbf{r}_1, \dots, \mathbf{r}_n)$  is in  $\mathcal{J}_n(\Lambda(d, N))$ , then we must have some  $1 \leq k_1 < k_2 \leq n$  such that  $\mathbf{r}_{k_1} = \mathbf{r}_{k_2}$ . Now we will make  $(\tilde{\mathbf{r}}_1(\ell), \dots, \tilde{\mathbf{r}}_n(\ell))$  for  $\ell \in \{0, \dots, \lfloor L/2 \rfloor - 1\}$ , as follows. If  $r_{k_1, 1} \in \{1, \dots, \lfloor L/2 \rfloor\}$  then we take  $\tilde{r}_{k_1, 1}(\ell) = r_{k_1, 1} + \ell$ , and we let  $\tilde{r}_{k, j}(\ell) = r_{k, j}$  for all  $(k, j) \neq (k_1, 1)$ . Otherwise, we have  $r_{k_1, 1} \in \{\lfloor L/2 \rfloor + 1, \dots, L\}$  and we let  $\tilde{r}_{k_1, 1}(\ell) = r_{k_1, 1} - \ell$ . We again have the same type of bound as in (52), except now we have  $\mathcal{J}_n(\Lambda(d, N))$ , instead of  $\mathbb{B}^d(L)^n \setminus \Lambda(d, N)^n$ .

We must again ask, for an arbitrary edge, going from  $(\mathbf{r}'_1, \dots, \mathbf{r}'_n)$  to  $(\mathbf{r}''_1, \dots, \mathbf{r}''_n)$ , how many times will it occur as an edge, from  $(\tilde{\mathbf{r}}_1(\ell; \mathbf{r}_1, \dots, \mathbf{r}_n))$  to  $(\tilde{\mathbf{r}}_1(\ell-1; \mathbf{r}_1, \dots, \mathbf{r}_n))$ , for some point  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathcal{J}_n(\Lambda(d, N))$ . Note that there is exactly one  $k_1 \in \{1, \dots, n\}$  such that  $\tilde{\mathbf{r}}'_{k_1} \neq \tilde{\mathbf{r}}''_{k_1}$ . There is also exactly one  $j \in \{1, \dots, d\}$  such that  $r'_{k_1, j} \neq r''_{k_1, j}$ . If  $j \neq 1$ , then the number of times the edge occurs is 0. But if  $j = 1$  then we may do the following. Consider  $\mathbf{r}'''(\ell) = (\ell, \mathbf{r}_{k_1, 2}, \dots, \mathbf{r}_{k_1, n})$ . Whenever we obtain that  $\mathbf{r}'''(\ell) = \mathbf{r}'_k$  for some  $k \neq k_1$  and  $\ell \in \{1, \dots, L\}$  then we may consider the  $n$ -tuple where, starting with  $(\mathbf{r}'_1, \dots, \mathbf{r}'_n)$ , we replace  $\mathbf{r}'_{k_1}$  by  $\mathbf{r}'''(\ell)$ . Then this is a possible starting point  $(\mathbf{r}_1, \dots, \mathbf{r}_n)$ . But these are the only possible starting points, and there are only  $n-1$  possibilities because we must have  $k \in \{1, \dots, n\} \setminus \{k_1\}$ .

Therefore, as before, we obtain

$$\sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathcal{J}_n(\Lambda(d, N))} |F(\mathbf{r}_1, \dots, \mathbf{r}_n)|^2 \leq \frac{2(n-1)L}{3} \langle F, \mathfrak{H}_{\mathbb{B}^d(L)}^{(n)} F \rangle + \frac{4n(n-1)d}{\lfloor L/2 \rfloor} \|F\|^2.$$

Putting this together with (53) gives the desired result.  $\square$

## 9 Proof of Key Step II – Part D: Extension theorem type bound

We will now prove Proposition 6.7. Suppose that we have  $F \in \ell^2(\mathbb{B}^d(L)^n)$  which is supported on  $\Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ . Let us define

$$\begin{aligned} & \tilde{\mathfrak{H}}_{\Lambda(d, N)}^{(n)} F(\mathbf{r}_1, \dots, \mathbf{r}_n) \\ &= \sum_{(\mathbf{r}'_1, \dots, \mathbf{r}'_n) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))} \mathbf{1}_{\{1\}}(\|(\mathbf{r}_1, \dots, \mathbf{r}_n) - (\mathbf{r}'_1, \dots, \mathbf{r}'_n)\|) [F(\mathbf{r}_1, \dots, \mathbf{r}_n) - F(\mathbf{r}'_1, \dots, \mathbf{r}'_n)]. \end{aligned} \quad (54)$$

Then the goal is to prove that there is an extension  $\tilde{F}$  of  $F$  to all of  $\mathbb{B}^d(L)^d$ , such that

- (i)  $\tilde{F}(\mathbf{r}_1, \dots, \mathbf{r}_n) = F(\mathbf{r}_1, \dots, \mathbf{r}_n)$  for each  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ ,
- (ii) there exists some fixed  $C_3(d, n)$  (not depending on  $F$ ) such that

$$\langle \tilde{F}, \tilde{\mathfrak{H}}_{\mathbb{B}^d(L)}^{(n)} \tilde{F} \rangle \leq C_3(d, n) \langle F, \tilde{\mathfrak{H}}_{\Lambda(d, N)}^{(n)} F \rangle. \quad (55)$$

Given any  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \notin \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ , let  $\rho(\mathbf{r}_1, \dots, \mathbf{r}_n)$  be the minimum distance to  $\Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ :

$$\rho(\mathbf{r}_1, \dots, \mathbf{r}_n) \stackrel{\text{def}}{=} \min\{\|(\mathbf{r}_1, \dots, \mathbf{r}_n) - (\mathbf{r}'_1, \dots, \mathbf{r}'_n)\|_1 : (\mathbf{r}'_1, \dots, \mathbf{r}'_n) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))\}.$$

Then, given  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n$  satisfying  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \notin \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ , we define  $(\tilde{\mathbf{r}}_1, \dots, \tilde{\mathbf{r}}_n) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$  to be the point such that  $\|(\mathbf{r}_1, \dots, \mathbf{r}_n) - (\tilde{\mathbf{r}}_1, \dots, \tilde{\mathbf{r}}_n)\|_1 = \rho(\mathbf{r}_1, \dots, \mathbf{r}_n)$  and such that  $(\tilde{\mathbf{r}}_1, \dots, \tilde{\mathbf{r}}_n)$  is minimal in the lexicographic ordering in case more than 1 such point exists. For later reference, let us denote these points as  $(\tilde{\mathbf{r}}_1(\mathbf{r}_1, \dots, \mathbf{r}_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}_1, \dots, \mathbf{r}_n))$ . Note that

$$(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N)) \quad \Leftrightarrow \quad (\tilde{\mathbf{r}}_1(\mathbf{r}_1, \dots, \mathbf{r}_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}_1, \dots, \mathbf{r}_n)) = (\mathbf{r}_1, \dots, \mathbf{r}_n).$$

Then we define, for every  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n$ ,

$$\tilde{F}(\mathbf{r}_1, \dots, \mathbf{r}_n) \stackrel{\text{def}}{=} F(\tilde{\mathbf{r}}_1(\mathbf{r}_1, \dots, \mathbf{r}_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}_1, \dots, \mathbf{r}_n)). \quad (56)$$

This does satisfy  $\tilde{F}(\mathbf{r}_1, \dots, \mathbf{r}_n) = F(\mathbf{r}_1, \dots, \mathbf{r}_n)$  if  $(\mathbf{r}_1, \dots, \mathbf{r}_n)$  is in  $\Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ , condition (i). So we just have to check condition (ii), namely equation (55).

Now, the key point is the following: suppose that  $(\mathbf{r}_1, \dots, \mathbf{r}_n)$  and  $(\mathbf{r}'_1, \dots, \mathbf{r}'_n)$  are neighbors in  $\mathbb{B}^d(L)^n$ , then we have

$$\begin{aligned} & \tilde{F}(\mathbf{r}_1, \dots, \mathbf{r}_n) - \tilde{F}(\mathbf{r}'_1, \dots, \mathbf{r}'_n) \\ &= F(\tilde{\mathbf{r}}_1(\mathbf{r}_1, \dots, \mathbf{r}_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}_1, \dots, \mathbf{r}_n)) - F(\tilde{\mathbf{r}}_1(\mathbf{r}'_1, \dots, \mathbf{r}'_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}'_1, \dots, \mathbf{r}'_n)). \end{aligned} \quad (57)$$

The question is how to bound the right hand side, using  $\langle F, \tilde{H}_{\Lambda(d, N)}^{(n)} F \rangle$ . We are somewhat motivated by the argument of Section 8. We will construct a chain connecting  $(\tilde{\mathbf{r}}_1(\mathbf{r}_1, \dots, \mathbf{r}_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}_1, \dots, \mathbf{r}_n))$  to  $(\tilde{\mathbf{r}}_1(\mathbf{r}'_1, \dots, \mathbf{r}'_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}'_1, \dots, \mathbf{r}'_n))$  within  $\Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$  and then we will use a telescoping sum.

We combine these two lemmas:

**Lemma 9.1** *There is a constant  $\rho_{\max}(n, d)$  such that for any  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)$ , we have*

$$\rho(\mathbf{r}_1, \dots, \mathbf{r}_n) \leq \rho_{\max}(n, d).$$

**Lemma 9.2** *There is a constant  $c_3(d, n)$  such that the following holds. If  $(\mathbf{r}_1, \dots, \mathbf{r}_n)$  and  $(\mathbf{r}'_1, \dots, \mathbf{r}'_n)$  are two points in  $\Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$  satisfying*

$$\|(\mathbf{r}_1, \dots, \mathbf{r}_n) - (\mathbf{r}'_1, \dots, \mathbf{r}'_n)\| \leq 2\rho_{\max}(n, d) + 1,$$

*then, for some  $\tau \leq c_3(d, n)$ , there are points  $(\hat{\mathbf{r}}_1(t), \dots, \hat{\mathbf{r}}_n(t)) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ , for  $t \in \{0, \dots, \tau\}$ , satisfying*

- $(\hat{\mathbf{r}}_1(0), \dots, \hat{\mathbf{r}}_n(0)) = (\mathbf{r}_1, \dots, \mathbf{r}_n),$
- $(\hat{\mathbf{r}}_1(\tau), \dots, \hat{\mathbf{r}}_n(\tau)) = (\mathbf{r}_{\pi_1}, \dots, \mathbf{r}_{\pi_n}),$  for some permutation  $\pi \in S_n$ , where  $\pi$  equals the identity if  $d > 1$ , and
- for each  $t \in \{1, \dots, \tau\},$

$$\|(\hat{\mathbf{r}}_1(t), \dots, \hat{\mathbf{r}}_n(t)) - (\hat{\mathbf{r}}_1(t-1), \dots, \hat{\mathbf{r}}_n(t-1))\|_1 = 1.$$

These lemmas are not difficult. But they require detailed descriptions. We wish to state their implication, first, which is the proof of Proposition 6.7.

**Proof of Proposition 6.7:** Suppose that  $(\mathbf{r}_1, \dots, \mathbf{r}_n)$  and  $(\mathbf{r}'_1, \dots, \mathbf{r}'_n)$  are two points in  $\mathbb{B}^d(L)^n$  satisfying

$$\|(\mathbf{r}_1, \dots, \mathbf{r}_n) - (\mathbf{r}'_1, \dots, \mathbf{r}'_n)\|_1 \leq 1.$$

Then by Lemma 9.1 and the triangle inequality, we see that

$$\|(\tilde{\mathbf{r}}_1(\mathbf{r}_1, \dots, \mathbf{r}_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}_1, \dots, \mathbf{r}_n)) - (\tilde{\mathbf{r}}_1(\mathbf{r}'_1, \dots, \mathbf{r}'_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}'_1, \dots, \mathbf{r}'_n))\|_1 \leq 2\rho_{\max}(n, d) + 1.$$

Then, by Lemma 9.2, there is a chain  $(\hat{\mathbf{r}}_1(t), \dots, \hat{\mathbf{r}}_n(t)) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ , for  $t \in \{0, \dots, \tau\}$  linking  $(\tilde{\mathbf{r}}_1(\mathbf{r}_1, \dots, \mathbf{r}_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}_1, \dots, \mathbf{r}_n))$  and  $(\tilde{\mathbf{r}}_{\pi_1}(\mathbf{r}'_1, \dots, \mathbf{r}'_n), \dots, \tilde{\mathbf{r}}_{\pi_n}(\mathbf{r}'_1, \dots, \mathbf{r}'_n))$ , for some  $\pi \in S_n$ . The fact that there is a permutation, does not concern us because  $F$  is a symmetric function. In other words, we then have, by (57) a telescoping sum

$$\tilde{F}(\mathbf{r}_1, \dots, \mathbf{r}_n) - \tilde{F}(\mathbf{r}'_1, \dots, \mathbf{r}'_n) \leq \sum_{t=1}^{\tau} F((\hat{\mathbf{r}}_1(t), \dots, \hat{\mathbf{r}}_n(t))) - F((\hat{\mathbf{r}}_1(t-1), \dots, \hat{\mathbf{r}}_n(t-1))).$$

Then, by the Cauchy-Schwarz inequality, we obtain

$$|\tilde{F}(\mathbf{r}_1, \dots, \mathbf{r}_n) - \tilde{F}(\mathbf{r}'_1, \dots, \mathbf{r}'_n)|^2 \leq \tau \sum_{t=1}^{\tau} |F((\hat{\mathbf{r}}_1(t), \dots, \hat{\mathbf{r}}_n(t))) - F((\hat{\mathbf{r}}_1(t-1), \dots, \hat{\mathbf{r}}_n(t-1)))|^2.$$

But by Lemma 9.2, we know that  $\tau \leq c_3(d, n)$ . So we obtain

$$|\tilde{F}(\mathbf{r}_1, \dots, \mathbf{r}_n) - \tilde{F}(\mathbf{r}'_1, \dots, \mathbf{r}'_n)|^2 \leq c_3(d, n) \sum_{t=1}^{\tau} |F((\hat{\mathbf{r}}_1(t), \dots, \hat{\mathbf{r}}_n(t))) - F((\hat{\mathbf{r}}_1(t-1), \dots, \hat{\mathbf{r}}_n(t-1)))|^2. \quad (58)$$

Now let us begin to enumerate how many times a given edge will be chosen in some term such as the right hand side above, when we apply this bound to every summand of the formula

$$\langle \tilde{F}, \mathfrak{H}_{\mathbb{B}^d(L)}^{(n)} \tilde{F} \rangle = \frac{1}{2} \sum_{\{(\mathbf{r}_1, \dots, \mathbf{r}_n), (\mathbf{r}'_1, \dots, \mathbf{r}'_n)\} \in \Phi_n(\mathbb{B}^d(L))} |\tilde{F}(\mathbf{r}_1, \dots, \mathbf{r}_n) - \tilde{F}(\mathbf{r}'_1, \dots, \mathbf{r}'_n)|^2, \quad (59)$$

We do not seek a reasonable bound, just a finite one. Therefore, given any pair

$$(\mathbf{r}_1^{(1)}, \dots, \mathbf{r}_n^{(1)}), (\mathbf{r}_1^{(2)}, \dots, \mathbf{r}_n^{(2)}) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N)),$$

of distance 1 apart, we see that if it is in a chain then there are at most

$$\mathfrak{N}(d, n) \stackrel{\text{def}}{=} \sum_{\tau=2}^{c_3(d, n)} (\tau - 1)(nd)^{\tau-1},$$

choices for how the chain continues to the left and the right (since at each step, the chain has at most  $(nd)$  choices for how to take its next step, and the length of the left and right sides of the chain emanating from our initial edge must be  $(k, \tau - 1 - k)$  for some  $k \in \{0, \dots, \tau - 1\}$ ). Given  $((\hat{\mathbf{r}}_1(0), \dots, \hat{\mathbf{r}}_n(0)))$  and  $((\hat{\mathbf{r}}_1(\tau), \dots, \hat{\mathbf{r}}_n(\tau)))$ , there is still a choice of the two points  $(\mathbf{r}_1, \dots, \mathbf{r}_n), (\mathbf{r}'_1, \dots, \mathbf{r}'_n) \in \mathbb{B}^d(L)^n$  such that

$$(\tilde{\mathbf{r}}_1(\mathbf{r}_1, \dots, \mathbf{r}_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}_1, \dots, \mathbf{r}_n)) = ((\hat{\mathbf{r}}_1(0), \dots, \hat{\mathbf{r}}_n(0)))$$

and

$$(\tilde{\mathbf{r}}_1(\mathbf{r}'_1, \dots, \mathbf{r}'_n), \dots, \tilde{\mathbf{r}}_n(\mathbf{r}'_1, \dots, \mathbf{r}'_n)) = ((\hat{\mathbf{r}}_1(\tau), \dots, \hat{\mathbf{r}}_n(\tau))).$$

But, no matter what value of  $\tau$  we have, we may just use Lemma 9.1 to say that  $\|(\mathbf{r}_1, \dots, \mathbf{r}_n) - ((\hat{\mathbf{r}}_1(0), \dots, \hat{\mathbf{r}}_n(0)))\|_1$  and  $\|(\mathbf{r}'_1, \dots, \mathbf{r}'_n) - ((\hat{\mathbf{r}}_1(\tau), \dots, \hat{\mathbf{r}}_n(\tau)))\|_1$  are bounded by  $\rho_{\max}(n, d)$ . Since  $((\hat{\mathbf{r}}_1(0), \dots, \hat{\mathbf{r}}_n(0)))$  and  $((\hat{\mathbf{r}}_1(\tau), \dots, \hat{\mathbf{r}}_n(\tau)))$  have been already chosen, this implies that there are at most  $(2\rho_{\max}(d, n) + 1)^{nd}$  choices for each of the points. But there is also a possibility of a permutation  $\pi \in S_n$ . So, we have determined that the pair of points  $(\mathbf{r}_1^{(1)}, \dots, \mathbf{r}_n^{(1)}), (\mathbf{r}_1^{(2)}, \dots, \mathbf{r}_n^{(2)}) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$  can come from at most

$$\mathfrak{N}'(d, n) \stackrel{\text{def}}{=} (2\rho_{\max}(d, n) + 1)^{2nd} (n!) \mathfrak{N}(d, n)$$

terms in the right hand side of (58) when we use these to bound all the summands in (59). Thus, combining all this, we obtain

$$\langle \tilde{F}, \mathfrak{H}_{\mathbb{B}^d(L)}^{(n)} \tilde{F} \rangle \leq \frac{1}{2} c_3(d, n) \mathfrak{N}'(d, n) \langle F, \tilde{H}_{\Lambda(d, N)}^{(n)} F \rangle.$$

□

One clear fact is that our bounds are quite large. This will also be the case in the proof of Lemma 9.1 and Lemma 9.2.

**Proof of Lemma 9.1:** Let us write  $\mathbf{R}$  for  $(\mathbf{r}_1, \dots, \mathbf{r}_n)$ , in order to simplify notation. Given  $\mathbf{R} = (\mathbf{r}_1, \dots, \mathbf{r}_n) \in \mathbb{B}^d(L)^n$  we will find a  $\tau$  and a sequence  $\check{\mathbf{R}}(t) = (\check{\mathbf{r}}_1(t), \dots, \check{\mathbf{r}}_n(t)) \in \mathbb{B}^d(L)^n$  for  $t \in \{0, \dots, \tau\}$ , which is a chain (so that  $\|\check{\mathbf{R}}(t) - \check{\mathbf{R}}(t-1)\|_1 = 1$  for each  $t \in \{1, \dots, \tau\}$ ) such that  $\check{\mathbf{R}}(0) = \mathbf{R}$  and  $\check{\mathbf{R}}(\tau) \in \Lambda(d, N)$ . Then if we find a uniform bound on  $\tau$ , that will give  $\rho_{\max}(d, n)$ .

Given  $\mathbf{R}$ , we will call this a point-vector. We will call each  $\mathbf{r}_1, \dots, \mathbf{r}_k$  the points. Sometimes we just refer to  $k$  as point index. We refer to  $j \in \{1, \dots, d\}$  as a coordinate index, since those index the coordinates  $\mathbf{r}_k = (r_{k,1}, \dots, r_{k,d})$  for some point  $\mathbf{r}_k$ .

**Step 1:** First let  $\tau_1 = \sum_{k=1}^n \sum_{j=1}^d \mathbf{1}_{\{L\}}(r_{k,j})$ . We want to push all coordinates of all points back from  $L$ , so that at the end no point of  $\check{\mathbf{R}}(\tau_1)$  is in  $\mathbb{B}^d(L)$ . I.e., we enumerate those pairs  $(k, j)$  such that  $r_{k,j} = L$ . At each  $t \in \{1, \dots, \tau_1\}$ , we choose the  $t$ th pair  $(k, j)$  and we let  $r_{k,j}(t) = L - 1$ , whereas, we had  $r_{k,j}(t-1) = L$ .

Note that we can only move 1 coordinate of 1 point at each step. So it suffices to say which one of these changes and how it changes (by going up or down by 1 step).

So, in the end,  $\check{\mathbf{R}}(\tau_1)$  has  $\check{r}_{k,j}(\tau_1) = \min\{L-1, r_{k,j}\}$  for each  $(k, j)$ . Note,  $\tau_1 \leq nd$ .

Step 2: This is the last step for this proof. It is harder. It involves a procedure we call “clustering” and “spreading apart.”

Let us define  $\approx$  to be a symmetric relation on  $\{1, \dots, n\}$  (although not transitive) wherein  $k_1 \approx_n k_2$  if  $\|\check{\mathbf{r}}_{k_1}(\tau) - \check{\mathbf{r}}_{k_2}(\tau)\|_1 < 2n-1$ . The reason for the choice of the distance,  $2n-1$ , will become more apparent, later. Then we define an equivalence relation  $\sim$  wherein  $k_1 \sim k_2$  if and only if there is some  $m$  and some sequence  $\hat{k}_0, \dots, \hat{k}_m$  such that  $\hat{k}_0 = k_1$ ,  $\hat{k}_m = k_2$  and  $\hat{k}_p \approx_n \hat{k}_{p-1}$  for each  $p \in \{1, \dots, m\}$ . We define the equivalence classes of  $\sim$  to be the “clusters.”

Now we will only focus on the first coordinate  $j = 1$ , in this proof. (But in the proof of Lemma 9.2, we will need an algorithm similar to this applied to all the coordinates.) Our goal is to find  $\tau_2$  and to extend  $\check{\mathbf{R}}(\tau_1 + t)$  for  $t \in \{1, \dots, \tau_2\}$  so that  $\check{r}_{1,1}(\tau_1 + \tau_2), \check{r}_{2,1}(\tau_1 + \tau_2), \dots, \check{r}_{n,1}(\tau_1 + \tau_2)$  are all distinct. I.e., we want the points of  $\check{\mathbf{R}}(\tau_1 + \tau_2)$  to have distinct 1st coordinates.

Suppose  $\mathcal{K}_1, \dots, \mathcal{K}_m$  are the clusters of  $\{1, \dots, n\}$ . For each  $p \in \{1, \dots, m\}$ , we enumerate the cluster  $\mathcal{K}_p$  as  $\{k_{p,1}, \dots, k_{p,|\mathcal{K}_p|}\}$  for some  $r(p)$ , where we choose the ordering such that

$$\check{r}_{k_{p,r},1}(\tau_1) \leq \check{r}_{k_{p,s},1}(\tau_1),$$

for each pair  $r, s$  with  $1 \leq r < s \leq |\mathcal{K}_p|$ . (In case the 1st coordinate of some point indices in  $\mathcal{K}_p$  coincide, we may choose the ordering for those coinciding point indices in any arbitrary way, such as the usual order.)

Now, if  $\check{r}_{k_{p,1},1}(\tau_1) \in \{\lfloor L/3 \rfloor, \dots, L-1\}$ , then we say that  $p$  enumerates a high cluster, or  $\mathcal{K}_p$  is a high cluster. If this fails to happen, but  $\check{r}_{k_{p,|\mathcal{K}_p|},1}(\tau_1) \in \{1, \dots, L - \lfloor L/3 \rfloor\}$ , then we say  $p$  enumerates a low cluster.

Now, also, note that since  $k_{p,1} \sim k_{p,|\mathcal{K}_p|}$ , this means that

$$k_{p,|\mathcal{K}_p|} - k_{p,1} \leq (|\mathcal{K}_p| - 1)(n - 1) \leq (n - 1)^2.$$

Therefore, if  $\mathcal{K}_p$  is not a high cluster, then  $k_{p,1} < \lfloor L/3 \rfloor$ , and that means  $k_{p,|\mathcal{K}_p|} \leq \lfloor L/3 \rfloor + (n-1)^2 - 1$ . So, in this case  $\mathcal{K}_p$  is a low cluster, as long as we have

$$L - \lfloor L/3 \rfloor \geq \lfloor L/3 \rfloor + (n-1)^2 - 1 \quad \Leftrightarrow \quad L - 2\lfloor L/3 \rfloor \geq (n-1)^2 - 1 = n(n-2)$$

In turn, this is insured if  $L \geq 3n(n-2)$ . So, henceforth, we assume  $L$  satisfies this lower bound. We also always assume  $L \geq 3n$  (in case  $n-2 \leq 1$ ).

High cluster case: Recursively, for each  $p \in \{1, \dots, m\}$ , we define  $\tau_{2,p}$  and  $\check{\mathbf{R}}(\tau_1 + \tau_{2,1} + \dots + \tau_{2,p-1} + t)$  for  $t \in \{1, \dots, \tau_{2,p}\}$  in order to do the following. If  $\mathcal{K}_p$  is a high cluster, then we take

$$\tau_{2,p} = \sum_{r=1}^{|\mathcal{K}_p|} \check{r}_{k_{p,r},1}(\tau_1 + \tau_{2,1} + \dots + \tau_{2,p-1}) - \sum_{r=1}^{|\mathcal{K}_p|} [\check{r}_{k_{p,r},1}(\tau_1 + \tau_{2,1} + \dots + \tau_{2,p-1}) + r - 1].$$

This is the minimum number of steps needed to send the coordinates  $\check{r}_{k_{p,r},1}(\tau_1 + \tau_{2,1} + \dots + \tau_{2,p-1})$  to  $[\check{r}_{k_{p,r},1}(\tau_1 + \tau_{2,1} + \dots + \tau_{2,p-1}) + r - 1]$ . That is how we prescribe  $\check{\mathbf{R}}(\tau_1 + \tau_{2,1} + \dots + \tau_{2,p-1} + t)$  for  $t \in \{1, \dots, \tau_{2,p}\}$ .

Actually, now we will be more precise about a particular order for the updates (because this will be useful in the proof of Lemma 9.2). Let us say that  $k_{p,r}$  is a “left-mover” if

$$\check{r}_{k_{p,r},1}(\tau_1 + \tau_{2,1} + \dots + \tau_{2,p-1}) > [\check{r}_{k_{p,r},1}(\tau_1 + \tau_{2,1} + \dots + \tau_{2,p-1}) + r - 1],$$

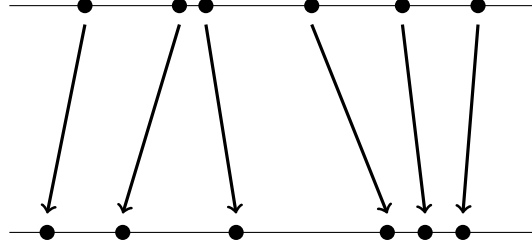


Figure 2: Left-movers are 1st, 2nd, 6th particles, and right-movers are the 3rd, 4th and 5th.

and say it is a “right-mover” if the strict inequality is reversed. We say it is stationary if there is equality instead of a strict inequality. See Figure 2.

We move the left-movers first, and then the right-movers, second. Moreover, we move the left-movers, starting the the left-most one, and proceeding to the right. Then since the final point coordinates are ordered (in the same order), we can see that no left mover collides with a right-mover which is staying still, since any right-mover to the left of the left-mover initially, also ends up to the left of the left-mover finally, and it starts even more to the left than that (while it is staying still). Also, no left mover collides with any right mover to its right, because the left-mover is moving right. By the time the left-mover moves, all the left-movers to its left have moved to their final positions, which is to the left of its final positions. So it does not collide with any of those, and it also does not collide with any of the left-movers to its right, because they stay stationary during its moves, and they are to its right (like the right-movers to the right). A similar argument applies to the right-movers when it is their turn to move. (One could do reflection and time-reversal to flip the roles of left- and right-movers, to see this, if one so desired.)

This focus on non-collision will be more important during the proof of Lemma 9.2.

Low cluster case: If  $\mathcal{K}_p$  is a low cluster, then it will be clear by symmetry what we do, based on the high cluster case. We take

$$\tau_{2,p} = - \sum_{r=1}^{|\mathcal{K}_p|} \check{r}_{k_p, |\mathcal{K}_p|+1-r, 1} (\tau_1 + \tau_{2,1} + \cdots + \tau_{2,p-1}) + \sum_{r=1}^{|\mathcal{K}_p|} \left[ \check{r}_{k_p, |\mathcal{K}_p|, 1} (\tau_1 + \tau_{2,1} + \cdots + \tau_{2,p-1}) + 1 - r \right].$$

This is the minimum number of steps needed to send the coordinates  $\check{r}_{k_p, |\mathcal{K}_p|+1-r, 1} (\tau_1 + \tau_{2,1} + \cdots + \tau_{2,p-1})$  to  $\left[ \check{r}_{k_p, |\mathcal{K}_p|, 1} (\tau_1 + \tau_{2,1} + \cdots + \tau_{2,p-1}) + 1 - r \right]$  for each  $r \in \{1, \dots, |\mathcal{K}_p|\}$ .

We do it in the same way as before, non-colliding. We will not repeat the details, here.

Conclusion: We know that the 1st coordinate of all the points in a given cluster are now spread apart, so they are not intersecting. We still need to check that points from distinct clusters did not move their first coordinates so as to now intersect. But in the high case, we moved the points at most  $|\mathcal{K}_p| - 1$  points to the right (for example if they all began intersecting at the coordinate of the left-most point). We did not move them to the left of the left-most point. So they are within  $n - 1$  of the original left-most-point. Similarly, in the low case, they are within  $n - 1$  of the right-most-point.

But, initially, points from distinct clusters have their first coordinates separated by at least  $2n - 1$ . So after the final time, we do not have any points from distinct clusters with 1st coordinate smaller than 1 apart. Thus we have achieved the desired goal of moving all the 1st coordinates apart at time  $\tau_1 + \tau_2$ , where  $\tau_2 = \tau_{2,1} + \cdots + \tau_{2,m}$ .

Now we want to bound  $\tau_2$ . Note that in each cluster, initially the diameter is at most  $(n - 1)(2n - 2)$ . So some thought shows that each point moves at most  $\max\{n - 1, (n - 1)(2n - 2)\}$ . Since there are  $n$  points to move, we can bound  $\tau_2$  by  $n \max\{n - 1, (n - 1)(2n - 2)\}$ . (We assume  $n > 1$ ,

otherwise there was no need for Step 2.) Since we already had a bound on  $\tau_1$ , we have a bound on  $\tau = \tau_1 + \tau_2$ . We call this bound  $\rho_{\max}(n, d)$ . It is  $nd + n \max\{n - 1, (n - 1)(2n - 2)\}$ .  $\square$

**Proof of Lemma 9.2:** We will construct a path as follows. We will construct 2 paths,  $\hat{\mathbf{R}}(t)$  and  $\hat{\mathbf{R}}'(t)$  starting from  $\mathbf{R}$  and  $\mathbf{R}'$ , respectively, ending in the same point. Then we may join them there, to get a chain that goes from  $\mathbf{R}$  to  $\mathbf{R}'$ .

To begin with, we are going to perform the same algorithm for the two point-vectors  $\mathbf{R}$  and  $\mathbf{R}'$ , independently of each other. I.e., initially, the algorithm for  $\mathbf{R}$  will be done independently of what  $\mathbf{R}'$  happens to be, and vice-versa. We call this initialization.

Initialization: The first part of the initialization step involves fixing for the fact that  $\Lambda(d, N)$  may not be a perfect box, since this will obstruct some later steps. There are two possibilities due to the construction of  $\Lambda(d, N)$ . Either no point of  $\Lambda(d, N)$  has the first coordinate equal to  $L$ . Or else, for every point, if we replace its  $d$ th coordinate by  $L$ , then that new point is in  $\mathbb{B}^d(L)$ . That is because of the lexicographic order we used.

In case 1, we perform the STEP2 algorithm from the proof of Lemma 9.2, first on coordinate 1. Note that on the first coordinate, every point already has that coordinate in  $\{1, \dots, L - 1\}$ , as it also was in the proof of Lemma 9.2. After this, we have coordinate 1 for all points distinct.

Then we do the STEP1 algorithm for all coordinates 2 to  $d$ . This moves all the coordinate away from  $L$ . We do not cause any collisions because all points had the first coordinate distinct. Then we do STEP2 from the proof of Lemma 9.2 for each coordinate 2 to  $d$ , in turn.

In case 2, we perform the STEP2 algorithm from the proof of Lemma 9.2 on coordinate  $d$ , modified to replace  $L$  by  $L + 1$  (because now the  $d$ th coordinate may be in  $\{1, \dots, L\}$  instead of  $\{1, \dots, L - 1\}$ ). Then we perform the STEP1 algorithm for each of the coordinates  $1, \dots, d - 1$ . This does not create any collisions because all points have distinct  $d$ th coordinate. Then we perform STEP2 algorithm for each of the coordinates  $1, \dots, d - 1$ , in turn. Then we perform the STEP1 algorithm for the  $d$ th coordinate, in order to reduce its range to  $\{1, \dots, L - 1\}$ . This does not create any intersections because for each  $j \in \{1, \dots, d - 1\}$  all points have distinct  $j$ th coordinate. Then we perform STEP2 again for coordinate  $d$ , to get all of the points to have distinct  $d$ th coordinate.

The initialization step is over. Let us say that our new points are  $\mathbf{R}^{(2)}$  and  $\mathbf{R}^{(3)}$ , in place of  $\mathbf{R}$  and  $\mathbf{R}'$ , respectively.

Lining up coordinate sets: We point out that there is some bound  $\text{bd}(d, n)$  for the number of steps we have made, which depends only on  $d$  and  $n$ . For example, we may see that we could take  $2d\mathfrak{N}''$ . This means that we have moved each coordinate of each point at most  $\text{bd}(d, n)$ . But initially, each coordinates of each point of  $\mathbf{R}$  is within  $2\rho_{\max}(d, n) + 1$  from the corresponding coordinate and point of  $\mathbf{R}'$ . So, now, what is true is that they are within  $2\rho_{\max}(d, n) + 1 + 2\text{bd}(d, n)$ . We call this  $\text{bd}^{(1)}(d, n)$ .

We now do the following. For each coordinate index  $j$ , we order the set of coordinates of the points of  $\mathbf{R}^{(2)}$  and the set of coordinate of the points of  $\mathbf{R}^{(3)}$ . Then we move the former set to the latter set using the left-mover/right-mover algorithm to avoid collisions. Note that each point moved at most  $\text{bf}^{(1)}(d, n)$  because those were the distances between the points initially, and ordering only makes distances decrease (of the corresponding points, in the order). This gives us a new  $\mathbf{R}^{(4)}$ . For each coordinate index  $j$ , the set of  $j$ th coordinates of the points of  $\mathbf{R}^{(4)}$  is the same set as the  $j$ th coordinates of the points of  $\mathbf{R}^{(3)}$ . But there may be a permutation, induced.

We note that we have now done at most  $nd\text{bd}^{(1)}(d, n)$  more steps. Therefore,

$$\|\mathbf{R}^{(4)} - \mathbf{R}^{(3)}\| \leq (nd + 1)\text{bd}^{(2)}(d, n) \stackrel{\text{def}}{=} \text{bd}^{(3)}(d, n). \quad (60)$$

Conclusion for  $d = 1$ : If  $d = 1$ , then we cannot necessarily get rid of the permutation. In this case, we terminate. We permute all the point indices of the second point and path by the inverse of



the permutation. Then this will result in  $\pi \mathbf{R}'$  going to  $\pi \mathbf{R}^{(4)}$  and we choose  $\pi$  so that  $\pi \mathbf{R}^{(4)} = \mathbf{R}^{(3)}$ . Then doing the steps to go from  $\mathbf{R}$  to  $\mathbf{R}^{(3)}$  and then undoing the steps of the other chain goes from  $\pi \mathbf{R}^{(4)}$  to  $\pi \mathbf{R}'$ . This is what was claimed.

Conclusion for  $d > 1$ :

Coordinate, by coordinate fix the permutation. If we are fixing the permutation in coordinate  $j$ , then this will cause collisions in coordinate  $j$ . But since all the other coordinates have all distinct points, this does not cause us to enter  $\mathcal{J}_n(\Lambda(d, N))$ . Draw the permutation and perform nearest neighbor transformations one at a time as they appear on the diagram, stretching the vertical direction if necessary to remove degeneracies. This will diminish distance at each stage because we are permuting to the correct order. After a permutation is done in coordinate  $j$ , we end again with all distinct points. Then move on to the next coordinate.  $\square$

## Appendices

### A Proof of advanced calculus fact for “new lows”

**Proof of Lemma 4.5:** For any  $\epsilon \in (0, 1)$ , we know that there exists  $n_0$  such that for  $n \geq n_0$ , we have

$$(1 - \epsilon)Cn^{-p} \leq t_n \leq (1 + \epsilon)Cn^{-p}.$$

We let  $n_1$  be defined as

$$n_1 = \min\{n \in \{n_0, n_0 + 1, \dots\} : (1 - \epsilon)Cn^{-p} < \min\{t_a, \dots, t_{n_0}\}\}.$$

Note that  $\min\{t_a, \dots, t_{n_0}\}$  is strictly positive, and  $(1 - \epsilon)Cn^{-p}$  converges to 0. So there is an  $n$  in  $\{n_0, n_0 + 1, \dots\}$  such that  $(1 - \epsilon)Cn^{-p} < \min\{t_a, \dots, t_{n_0}\}$ .

Then we know that for all  $n \geq n_1$  we have

$$\min\{t_a, \dots, t_n\} \geq (1 - \epsilon)Cn^{-p}.$$

On the other hand we know that for any  $m \in \{n_0, n_0 + 1, \dots\}$ ,

$$t_m \leq (1 + \epsilon)Cm^{-p}.$$

So defining  $A(\epsilon; n) = \min\{k \in \mathbb{Z} : k \geq [(1 + \epsilon)/(1 - \epsilon)]^{1/p} n\}$ , which is an integer bigger than 1, we know that for  $n \in \{n_1, n_1 + 1, \dots\}$  and  $m \in \{A(\epsilon; n), A(\epsilon; n) + 1, \dots\}$ ,

$$t_m \leq \min\{t_a, \dots, t_n\}.$$

Now let  $k^* = \min\{k \in \{a, a + 1, \dots\} : t_k \leq t_{A(\epsilon; n)}\}$ . Note that with this choice we have  $k^* \leq A(\epsilon; n)$ . But by the last displayed equation for  $m = A(\epsilon; n)$  we also have  $k^* \geq n$ . Moreover, for any  $k < k^*$  we have  $t_k > t_{A(\epsilon; n)} \geq t_{k^*}$ . So  $t_{k^*} = \min\{t_a, \dots, t_{k^*}\}$ . So  $k^*$  is a “new low” in the interval  $\{n, \dots, A(\epsilon; n)\}$ .

So  $\mathcal{N}_{\text{low}}(n) \leq A(\epsilon; n)$  since this is the smallest new low in  $\{n, n + 1, \dots\}$ .

By hypothesis, we know that  $t_{\mathcal{N}_{\text{low}}(n)} \leq t_n \leq (1 + \epsilon)Cn^{-p}$ . But now we also know

$$t_{\mathcal{N}_{\text{low}}(n)} \geq C(1 - \epsilon)[\mathcal{N}_{\text{low}}(n)]^{-p} \geq C(1 - \epsilon)[A(\epsilon; n)]^{-p}.$$

Hence we see that

$$(1 - \epsilon)^2(1 + \epsilon)^{-1} = (1 - \epsilon) \liminf_{n \rightarrow \infty} [A(\epsilon; n)/n]^{-p} \leq \liminf_{n \rightarrow \infty} \frac{t_{\mathcal{N}_{\text{low}}(n)}}{Cn^{-p}} \leq \limsup_{n \rightarrow \infty} \frac{t_{\mathcal{N}_{\text{low}}(n)}}{Cn^{-p}} \leq 1 + \epsilon.$$

Since  $\epsilon \in (0, 1)$  is arbitrary, and the bracketing quantities both converge to 1 as  $\epsilon \rightarrow 0$ , this proves the claim.  $\square$

## B Proof of summary of spectrum of $\mathfrak{H}_{\mathbb{B}^d(L)}^{(n)}$

The full spectrum of  $\mathfrak{H}_{\mathbb{B}^d(L)}^{(n)}$  may be calculated as products of functions  $\prod_{k=1}^n \prod_{j=1}^d f(L^{-1}\kappa_{k,j}, r_{k,j})$ , unsymmetrized. This may be seen in  $d = 1$ . But then it follows from separation of variables for higher  $d$ . Then symmetrization is what is necessary to restrict to the range of  $\mathfrak{S}_{\mathbb{B}^d(L)}^{(n)}$ .

## C Proofs of easy facts about the graph Laplacians

We leave these lemmas as an exercise. They can be done using the fact that  $H_{\mathcal{G}} \upharpoonright \mathcal{H}_{\mathcal{V}}^{(n)}$  is unitarily equivalent to  $\tilde{\mathfrak{H}}_{\mathcal{G}}^{(n)}$  restricted to the range of  $\mathfrak{S}_{\mathcal{V}}^{(n)}$ , and properties of graph Laplacians: such as monotonicity relative to the psd cone under the action of adding edges.

## D Proof of Proposition 3.9

Based on the results of Appendix C, what we must prove is the following. Suppose that we have  $(\kappa_1, \dots, \kappa_n) \in (\{0, 1, \dots\}^d)^n$ . For some  $L_0$ , we have  $(\kappa_1, \dots, \kappa_n) \in (\{0, \dots, L_0 - 1\}^d)^n$ . Then we consider only values of  $L$  with  $L \geq L_0$ . We define  $\nu_{(\kappa_1, \dots, \kappa_n)} \in \mathcal{O}_n(d, L)$ . Then, considering  $\tilde{F}_{d,L}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \cdot)$ , we know by Lemma 7.3 that

$$\mathfrak{H}_{\mathbb{B}^d(L)}^{(n)} \tilde{F}_{d,L}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \cdot) = \lambda_{d,L}^{(n)}(\kappa_1, \dots, \kappa_n) \tilde{F}_{d,L}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \cdot),$$

where

$$\lambda_{d,L}^{(n)}(\kappa_1, \dots, \kappa_n) = \sum_{k=1}^n \sum_{j=1}^d 2 \sin^2 \left( \frac{\pi \kappa_{k,j}}{2L} \right). \quad (61)$$

Therefore, for any  $(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))$ , we have

$$\begin{aligned} \left( \tilde{\mathfrak{H}}_{\Lambda(d, N)}^{(n)} - \lambda_{d,L}^{(n)}(\kappa_1, \dots, \kappa_n) \right) \tilde{F}_{d,L}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \mathbf{r}_1, \dots, \mathbf{r}_n) \\ = \left( \tilde{\mathfrak{H}}_{\Lambda(d, N)}^{(n)} - \mathfrak{H}_{\mathbb{B}^d(L)}^{(n)} \right) \tilde{F}_{d,L}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \mathbf{r}_1, \dots, \mathbf{r}_n). \end{aligned}$$

Since everything is symmetric, this implies that

$$\begin{aligned} \left\| \left( H_{\Lambda(d, N)} - \lambda_{d,L}^{(n)}(\kappa_1, \dots, \kappa_n) \right) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n) \right\|^2 \\ = \sum_{(\mathbf{r}_1, \dots, \mathbf{r}_n) \in \Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))} \left| \left( \tilde{\mathfrak{H}}_{\Lambda(d, N)}^{(n)} - \mathfrak{H}_{\mathbb{B}^d(L)}^{(n)} \right) \tilde{F}_{d,L}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \mathbf{r}_1, \dots, \mathbf{r}_n) \right|^2. \end{aligned}$$

We may also rewrite this formula as

$$\begin{aligned} \left( \tilde{\mathfrak{H}}_{\Lambda(d, N)}^{(n)} - \mathfrak{H}_{\mathbb{B}^d(L)}^{(n)} \right) \tilde{F}_{d,L}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \mathbf{r}_1, \dots, \mathbf{r}_n) \\ = \frac{1}{2} \sum_{\substack{(\mathbf{r}'_1, \dots, \mathbf{r}'_n) \\ \|(\mathbf{r}'_1, \dots, \mathbf{r}'_n) - (\mathbf{r}_1, \dots, \mathbf{r}_n)\| = 1}} (1 - \mathbf{1}_{\Lambda(d, N)^n \setminus \mathcal{J}_n(\Lambda(d, N))}(\mathbf{r}'_1, \dots, \mathbf{r}'_n)) (F(\mathbf{r}_1, \dots, \mathbf{r}_n) - F(\mathbf{r}'_1, \dots, \mathbf{r}'_n)), \end{aligned}$$

where we have written  $F$  for  $\tilde{F}_{d,L}^{(n)}(\nu_{(\kappa_1, \dots, \kappa_n)}; \cdot)$  in this formula to reduce notation. The number of terms in this sum is always bounded by  $2nd$  (the degree of  $(\mathbb{Z}^d)^n \cong \mathbb{Z}^{nd}$ ). Therefore, by Cauchy-Schwarz, and a sup-norm bound, we may bound

$$\begin{aligned} \left\| \left( H_{\Lambda(d,N)} - \lambda_{d,L}^{(n)}(\kappa_1, \dots, \kappa_n) \right) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n) \right\|^2 \\ \leq nd \max_{\substack{(\mathbf{r}_1, \dots, \mathbf{r}_n), (\mathbf{r}'_1, \dots, \mathbf{r}'_n) \in (\mathbb{B}^d(L))^n \\ \|\mathbf{r}'_1, \dots, \mathbf{r}'_n - \mathbf{r}_1, \dots, \mathbf{r}_n\| = 1}} |F(\mathbf{r}_1, \dots, \mathbf{r}_n) - F(\mathbf{r}'_1, \dots, \mathbf{r}'_n)|^2 \cdot \mathcal{N}, \end{aligned}$$

where  $\mathcal{N}$  equals the number of edges in  $\Phi_n(\mathbb{B}^d(L)) \setminus \Theta_n(\Lambda(d, N))$ . For any pair  $(\mathbf{r}_1, \dots, \mathbf{r}_n), (\mathbf{r}'_1, \dots, \mathbf{r}'_n) \in (\mathbb{B}^d(L))^n$  with  $\|\mathbf{r}'_1, \dots, \mathbf{r}'_n - \mathbf{r}_1, \dots, \mathbf{r}_n\| = 1$ , we note, by the fundamental theorem of calculus (and Cauchy-Schwarz), that

$$|F(\mathbf{r}_1, \dots, \mathbf{r}_n) - F(\mathbf{r}'_1, \dots, \mathbf{r}'_n)|^2 \leq \int_0^1 \left\| \frac{d}{dt} F(t\mathbf{r}_1 + (1-t)\mathbf{r}'_1, \dots, t\mathbf{r}_n + (1-t)\mathbf{r}'_n) \right\|^2 dt,$$

where we exted the definition of  $F$  from  $(\mathbb{B}^d(L))^n$  to  $\mathbb{R}^{dn}$  by the same formula as in Definition 7.2. We note that only one coordinate  $k \in \{1, \dots, n\}$  has  $\mathbf{r}_k \neq \mathbf{r}'_k$ , and in this case there is only one  $j \in \{1, \dots, d\}$  such that  $r_{kj} \neq r'_{kj}$ . Then direct calculation shows

$$\left\| \frac{d}{dt} F(t\mathbf{r}_1 + (1-t)\mathbf{r}'_1, \dots, t\mathbf{r}_n + (1-t)\mathbf{r}'_n) \right\|^2 \leq L^{-nd} |\mathfrak{K}_{d,L}(\nu)|^{-1} 2^{nd} \cdot \frac{\pi^2 \kappa_{kj}^2}{L^2}.$$

So, what we see is that

$$\left\| \left( H_{\Lambda(d,N)} - \lambda_{d,L}^{(n)}(\kappa_1, \dots, \kappa_n) \right) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n) \right\|^2 \leq nd 2^{nd} \max_{\substack{k \in \{1, \dots, n\} \\ j \in \{1, \dots, d\}}} \frac{\kappa_{kj}^2}{L^2} \cdot \frac{\mathcal{N}}{L^{nd}}.$$

But we claim that there is some constant  $C(n, d)$  such that

$$\frac{\mathcal{N}}{L^{nd}} \leq \frac{C(n, d)}{L},$$

because at least one of the points  $(\mathbf{r}_1, \dots, \mathbf{r}_n), (\mathbf{r}'_1, \dots, \mathbf{r}'_n) \in (\mathbb{B}^d(L))^n$  must be in  $(\mathbb{B}^d(L))^n \setminus \mathcal{I}_n(\Lambda(d, N))$ . Assume it is  $(\mathbf{r}'_1, \dots, \mathbf{r}'_n)$ . Then this means that at least one of the two scenarios must occur. The first possibility is that for some  $k \in \{1, \dots, d\}$  we have  $\mathbf{r}'_k \in \mathbb{B}^d(L) \setminus \Lambda(d, N)$ , meaning that for some  $j \in \{1, \dots, d\}$  we must have  $r'_{kj} = L$ . This restricts that choice of that coordinate. Or, the second possibility is that for some pair  $1 \leq k_1 < k_2 \leq n$  we have  $\mathbf{r}'_{k_1} = \mathbf{r}'_{k_2}$ . This restricts all  $j$  coordinates  $r'_{k_2, j}$  once we have chosen  $\mathbf{r}'_{k_1}$ . So this proves the claim. Putting this all together, we obtain

$$\left\| \left( H_{\Lambda(d,N)} - \lambda_{d,L}^{(n)}(\kappa_1, \dots, \kappa_n) \right) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n) \right\|^2 \leq C'(d, n) \max_{\substack{k \in \{1, \dots, n\} \\ j \in \{1, \dots, d\}}} \frac{\kappa_{kj}^2}{L^3}.$$

Multiplying through by  $\gamma^{-1} \cdot L^2$ , we obtain

$$\left\| \left( \gamma^{-1} \cdot L^2 H_{\Lambda(d,N)} - \gamma^{-1} \cdot L^2 \lambda_{d,L}^{(n)}(\kappa_1, \dots, \kappa_n) \right) \tilde{\Psi}_{d,N}^{(n)}(\kappa_1, \dots, \kappa_n) \right\|^2 \leq C'(d, n) \max_{\substack{k \in \{1, \dots, n\} \\ j \in \{1, \dots, d\}}} \frac{\kappa_{kj}^2}{L}.$$

The right hand side does converge to 0 as  $L \rightarrow \infty$ . Moreover, by (61), we do have the simple formula

$$\lim_{L \rightarrow \infty} L^2 \lambda_{d,L}^{(n)}(\kappa_1, \dots, \kappa_n) = \lim_{L \rightarrow \infty} \sum_{k=1}^n \sum_{j=1}^d 2L^2 \sin^2 \left( \frac{\pi \kappa_{k,j}}{2L} \right) = \sum_{k=1}^n \sum_{j=1}^d \frac{\pi^2}{2} \kappa_{k,j}^2 = \gamma \sum_{k=1}^n \sum_{j=1}^d \kappa_{k,j}^2.$$

This proves (20).

In principle it is easier to prove (21) since one is operating at the level of  $\ell^2$  not energy. In fact, the main difference is that there will not be any need for the fundamental theorem of calculus. Everything follows from sup-norm bounds on the functions  $f$  (which is akin to a type of near equipartition of the  $\ell^2$  energy) along with the fact that the relative fraction of bad set of vertices is still bounded by  $\mathcal{N}/L^d = O(L^{-1})$ . We leave the details to the reader.

## Acknowledgments

This research was partially supported by the National Science Foundation under Grant DMS-1515850 (B.N.).

## References

- [1] G. Alon and G. Kozma. The probability of long cycles in interchange processes. *Duke Math. J.*, **162**, no. 9 (2013), 1567–1585.
- [2] G. Alon and G. Kozma. Ordering the Representations of Using the Interchange Process. *Canad. Math. Bull.* **56** (2013), 13–30.
- [3] Felix Bloch. *Zeitschrift für Physik*, **61**, 206–219 (1930) and ibidem **74** 295–?? (1932).
- [4] P. Caputo. Energy gap estimates in XXZ ferromagnets and stochastic particle systems. *Markov Proc. Rel. Fields* **11** (2005), 189–210.
- [5] P. Caputo, T. M. Liggett and T. Richthammer. Proof of Aldous’ spectral gap conjecture. *J. Amer. Math. Soc.* **23** (2010), 831–851.
- [6] F. Cesi. On the eigenvalues of Cayley graphs on the symmetric group generated by a complete multipartite set of transpositions, *J. Algebr. Combin.* **32**, no. 2 (2010), 155–185.
- [7] Joseph G. Conlon and Jan Philip Solovej. On asymptotic limits for the quantum Heisenberg model. *Journ. Phys. A: Math. Gen.* **23**, 3199–3213 (1990).
- [8] Joseph G. Conlon and Jan Philip Solovej. Random walk representations of the Heisenberg model. *Journ. Stat. Phys.* **64**, 251–270 (1991).
- [9] Joseph G. Conlon and Jan Philip Solovej. Upper bound on the free energy of the spin 1/2 Heisenberg ferromagnet. *Lett. Math. Phys.* **23**, 223–231 (1991).
- [10] M. Conomos and S. Starr. Asymptotics of the Spectral Gap for the Interchange Process on Large Hypercubes. *J. Stat. Mech.* (2011), P10018.
- [11] M. Correggi, A. Giuliani and R. Seiringer. Validity of spin-wave theory for the quantum Heisenberg model. *Europhys. Lett.* **108**, no. 2 (2014), 20003.

- [12] M. Correggi, A. Giuliani and R. Seiringer. Validity of the Spin-Wave Approximation for the Free Energy of the Heisenberg Ferromagnet. *Commun. Math. Phys.* **339**, no. 1, (2015), 279–307.
- [13] A. B. Dieker. Interlacings for Random Walks on Weighted Graphs and the Interchange Process. *SIAM J. Discrete Math.* **24**, no. 1 (2010), 191–206.
- [14] Freeman J. Dyson. General Theory of Spin-Wave Interactions. *Phys. Rev.* **102**, 1217–1230 (1956).
- [15] Freeman J. Dyson. Thermodynamic Behavior of an Ideal Ferromagnet. *Phys. Rev.* **102**, 1230–1244 (1956).
- [16] A. R. Edmonds. *Angular Momentum in Quantum Mechanics*. Princeton University Press, Princeton, NY, 1957.
- [17] L.C. Evans. *Partial Differential Equations. Graduate Studies in Mathematics, vol. 19*. American Mathematical Society, Providence, RI, 1998.
- [18] S. Handjani and D. Jungreis. Rate of Convergence for Shuffling Cards by Transpositions. *J. Theor. Prob.* **9**, pp. 983–993 (1996).
- [19] T. Koma and B. Nachtergaele. The Spectral Gap of the Ferromagnetic XXZ Chain. *Lett. Math. Phys.* **40**, pp. 1–16 (1997).
- [20] E.H. Lieb and D.C. Mattis. Ordering Energy Levels of Interacting Spin Systems. *J. Math. Phys.*, **3**, pp. 749–751 (1962).
- [21] Thomas M. Liggett. *Interacting Particle Systems. Grundlehren der mathematischen Wissenschaften 276*. Springer-Verlag, New York, 1985.
- [22] B. Morris. Spectral gap for the interchange process in a box. *Electr. Commun. Probab.* **13** (2008).
- [23] B. Nachtergaele, S. Ng and S. Starr. Ferromagnetic Ordering of Energy Levels for  $U_q(\mathfrak{sl}_2)$  Symmetric Spin Chains, *Lett. Math. Phys.* **100**, no. 3 (2012), 327–356.
- [24] B. Nachtergaele, W. Spitzer, and S. Starr. Ferromagnetic Ordering of Energy Levels. *J. Statist. Phys.* **116**, pp. 719–738 (2004).
- [25] B. Nachtergaele, W. Spitzer, and S. Starr. Droplet Excitations for the Spin-1/2 XXZ Chain with Kink Boundary Conditions, *Ann. Henri Poincaré*, **8**, no. 1 (2007), 165–201.
- [26] B. Nachtergaele and S. Starr. Ferromagnetic Lieb-Mattis Theorem. *Phys. Rev. Lett.* **94** (2005), 057206.
- [27] W. Spitzer, S. Starr, and L. Tran. Counterexamples to Ferromagnetic Ordering of Energy Levels. *J. Math. Phys.* **53** 043302 (2012).
- [28] H. N. V. Temperley and E. H. Lieb. Relations between the percolation and colouring problem and other graph-theoretical problems associated with regular planar lattices: some exact results for the percolation problem, *Proc. Roy. Soc.* **A322** (1971), 252–280.
- [29] Lawrence E. Thomas. Quantum Heisenberg ferromagnets and stochastic exclusion processes. *Journ. Math. Phys.* **21** (7), 1921–1924 (1980).

- [30] Balint Toth. Improved lower bound on the thermodynamic pressure of the spin  $1/2$  Heisenberg ferromagnet. *Lett. Math. Phys.* **28**, 75–84 (1993).